

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 42.7151 Seconds
(without alignments)
228.984 Million cell updates/sec

Title: US-09-700-643A-1
Perfect score: 167
Sequence: 1 SRAHQSMETPTDINPAWYAGRGIRPVGRX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL.25.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_plant.*
 - 10: sp_rodent.*
 - 11: sp_virus.*
 - 12: sp_vertebrate.*
 - 13: sp_unclassified.*
 - 14: sp_rv.*
 - 15: sp_rv.*
 - 16: sp_bacteriophage.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	166	99.4	98	Q8WN12	Q8WN12 ovis aries
2	98	58.7	117	Q9W624	Q9W624 carassius a
3	62	37.1	692	Q9L8J6	Q9L8J6 rhodospirill
4	58	34.7	1245	Q7X3H6	Q7X3H6 thiocapsa r
5	57	34.1	54	Q7Z6Y1	Q7Z6Y1 homo sapien
6	57	34.1	355	Q95VJ8	Q95VJ8 ciona savi
7	57	34.1	465	Q60687	Q60687 homo sapien
8	57	34.1	465	Q8WN85	Q8WN85 homo sapien
9	56.5	33.8	303	Q8YWC7	Q8YWC7 anabaena sp
10	56	33.5	625	Q89VA3	Q89VA3 bradyrhizob
11	55	32.9	1236	Q9JPA4	Q9JPA4 rhodocyclu
12	54.5	32.6	503	Q9KRY1	Q9KRY1 vibrio chol
13	54	32.3	1120	Q7UPG7	Q7UPG7 rhodospirell
14	53.5	32.0	176	Q08689	Q08689 saccharomyc
15	53	31.7	376	Q9RFP2	Q9RFP2 deinococcus
16	53	31.7	467	Q8R054	Q8R054 mus musculu

17	53	31.7	468	11	Q8K1F8	Q8K1F8 mus musculu
18	52	31.1	73	16	Q8U515	Q8U515 agrobacteri
19	52	31.1	308	16	Q8UC89	Q8UC89 agrobacteri
20	52	31.1	314	16	Q7VSA1	Q7VSA1 bordetella
21	52	31.1	419	16	Q8UKP2	Q8UKP2 agrobacteri
22	52	31.1	1253	2	Q8RTV7	Q8RTV7 uncultured
23	52	31.1	1252	2	Q9ZGE5	Q9ZGE5 heliobacill
24	51.5	30.8	130	16	Q82RV0	Q82RV0 streptomyce
25	51.5	30.8	327	10	Q94KU9	Q94KU9 brassica ca
26	51.5	30.8	327	10	Q94FZ9	Q94FZ9 brassica ca
27	51.5	30.8	480	16	Q7V5F5	Q7V5F5 prochloroco
28	51.5	30.8	664	16	Q911W4	Q911W4 pseudomonas
29	51	30.5	184	16	Q8PER0	Q8PER0 xanthomonas
30	51	30.5	302	5	Q19530	Q19530 caenorhabdi
31	51	30.5	315	2	Q87474	Q87474 burkholderi
32	51	30.5	333	4	Q968D4	Q968D4 homo sapien
33	51	30.5	367	4	Q8N6Q2	Q8N6Q2 homo sapien
34	51	30.5	367	4	Q81XT2	Q81XT2 homo sapien
35	51	30.5	370	11	Q8CGW9	Q8CGW9 mus musculu
36	51	30.5	430	17	Q27142	Q27142 methanobact
37	51	30.5	514	5	Q9VRV3	Q9VRV3 drosophila
38	51	30.5	689	11	Q8R3R1	Q8R3R1 mus musculu
39	51	30.5	767	11	Q8JZN2	Q8JZN2 mus musculu
40	51	30.5	798	11	Q91W50	Q91W50 mus musculu
41	51	30.5	822	11	Q80TP8	Q80TP8 mus musculu
42	51	30.5	1543	5	Q8T4U2	Q8T4U2 manduca sex
43	51	30.5	1845	17	Q8RTS7	Q8RTS7 methanosarc
44	50.5	30.2	292	10	Q8S7E3	Q8S7E3 oryza sativ
45	50.5	30.2	292	10	Q7XC01	Q7XC01 oryza sativ

ALIGNMENTS

RESULT 1

ID	Q8WN12	PRELIMINARY;	PRT;	98 AA.
AC	Q8WN12;			
DI	01-MAR-2002 (Tremblrel. 20, Created)			
DI	01-MAR-2002 (Tremblrel. 20, Last sequence update)			
DI	01-MAR-2002 (Tremblrel. 20, Last annotation update)			
DE	Preprolactin-releasing peptide.			
OS	Ovis aries (Sheep).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Caprinae; Ovis.			
OX	NCBI_taxid=9940;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Curler J.D., Kusters D.H.L., Barclay J.L., Anderson S.T.;			
RT	"Prolactin-releasing peptide (prp) in the ewe: cDNA cloning, mRNA			
RT	distribution and effects on prolactin secretion in vitro and in			
RT	vivo."			
RL	Submitted (NOV-2001) to the ENBL/GenBank/DBJ databases.			
DR	EMBL; AF450453; AAL47178.1; -			
SQ	SEQUENCE 98 AA; 10513 MW; 2A53331ED62CAAB5 CRC64;			

Query Match 99.4%; Score 166; DB 6; Length 98;

Best Local Similarity 100.0%; Pred. No. 2.1e-16;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQSMETPTDINPAWYAGRGIRPVGR 30

|||||

Db 23 SRAHQSMETPTDINPAWYAGRGIRPVGR 52

RESULT 2

ID	Q9W624	PRELIMINARY;	PRT;	117 AA.
AC	Q9W624;			
DI	01-NOV-1999 (Tremblrel. 12, Created)			
DI	01-NOV-1999 (Tremblrel. 12, Last sequence update)			
DI	01-DEC-2001 (Tremblrel. 19, Last annotation update)			


```
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.  
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 4.  
KW Metal-binding; Zinc; Zinc-finger;  
SQ SEQUENCE 355 AA; 40876 MW; B58F5DEDD812E9AC CRC64;  
  
Query Match 34.1%; Score 57; DB 5; Length 355;  
Best Local Similarity 40.0%; Pred.No. 6.1;  
Matches 12; Conservative 7; Mismatches 7; Indels 4; Gaps  
1;
```

```
QY 2 RAHQHSMEIRTPDIN----PAWYAGRGIRP 27  
      :|::||::: . | :|::||: |  
Db    27 QAQHSDSKPMQLNSVPSAAAYGYGMIP 56
```

```
RESULT 7  
ID O60687 PRELIMINARY; PRT; 465 AA.  
AC Q60687;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DI 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Sushi-repeat protein (Sushi-repeat containing protein).  
GN SRPLU.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kurosawa H., Inukai T., Inaba T., Goi K., Chang K.-S., Sinjyo T.,  
R Rakestraw K.M., Naeve C.W., Look T.A.;  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
[2]  
RP SEQUENCE FROM N.A.  
RA Huang C.-H., Chen H., Peng J., Chen Y.;  
RT "Cloning and characterization of the sushi-repeat containing protein  
RT (SRP) as a novel interaction partner of Rh type C glycoprotein  
RT (RhCG)." ;  
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF060567; AAC15765.1; -  
DR ENEL; AF393649; AAM73693.1; -  
GO; GO:006118; P:electron transport; IEA.  
DR InterPro; IPR001128; Cytochrome_P450.  
DR InterPro; IPR003410; HyalIn.  
DR InterPro; IPR000436; Sushi_SCR_CCP.  
PFam; PF02494; HVR; 1.  
PFam; PF00084; sushi; 3.  
DR PROSITE; PS00086; CYTOCHROME_P450_1.  
SQ SEQUENCE 465 AA; 52971 MW; 4D752B187FF3EFB8 CRC64;  
  
Query Match 34.1%; Score 57; DB 4; Length 465;  
Best Local Similarity 56.2%; Pred.No. 8.1;  
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps  
0;
```

```
QY 12 TPDPNPAWYAGRGIRP 27  
      ||:|::|:  
Db 18 TPAVTPTWYAGGYYP 33
```

```
RESULT 8  
Q8WW85 PRELIMINARY; PRT; 465 AA.  
ID Q8WW85  
AC Q8WW85;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DI 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Sushi-repeat protein.  
GN Sushi-repeat (Human).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.
```

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae; Bradyrhizobium.

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=USDA 110;
 RX MEDLINE=22484996; PubMed=12597275;
 RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
 RA Sasanoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
 RA Kohara M., Matsumoto M., Shimo S., Tezuka H., Wada T., Yamada M.,
 RA Tabata S.;
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
 RT Bradyrhizobium japonicum USDA110.";
 RL DNA Res. 9:189-197(2002).
 DR ENBL; AP005939; BAC46409.1; -;
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0003910; F:DNA ligase (ATP) activity; IEA.
 DR GO; GO:0006310; P:DNA recombination; IEA.
 DR GO; GO:0006281; P:DNA repair; IEA.
 DR GO; GO:0006260; P:DNA replication; IEA.
 DR InterPro; IPR000977; DNA_ligase.
 DR Pfam; PF01088; DNA_ligase; 1.
 DR Pfam; PF04679; DNA_ligase A.C; 1.
 DR PROSITE; PS00697; DNA_LIGASE_A1; 1.
 DR PROSITE; PS0160; DNA_LIGASE_A3; 1.
 KW Complete proteome.
 SQ SEQUENCE 625 AA; 69415 MW; 56BED3A80A46AD42 CRC64;

Query Match 33.5%; Score 56; DB 16; Length 625;
 Best Local Similarity 38.7%; Pred. No. 15;
 Matches 12; Conservative 4; Mismatches 9; Indels 6; Gaps 1;

OY 6 HSMETPTDNP-----AWYAGGIRPVGR 30
 |||||
 DB 240 HEVELWPGLAPPYLDLFWLEGRGKPVNR 270
 |||||

RESULT 11

ID O9JPA4 PRELIMINARY; PRT; 1236 AA.
 AC O9JPA4;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Mg protoporphyrin methyl transferase.
 GN BCHH.
 OS Rhodocyclus gelatinosus (Rhodopseudomonas gelatinosa).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Comamonadaceae; Rubrivivax.
 OX NCBI_TaxID=28068;
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL144;
 RX MEDLINE=94132007; PubMed=8300574;
 RA Nagashima K.V., Matsuura K., Ohnaya S., Shimada K.;
 RT "Phylogenetic analysis of photosynthetic genes of Rhodocyclus
 RT gelatinosus: Possibility of horizontal gene transfer in purple
 RT bacteria.";
 RL Photosyn. Res. 36:185-191(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL144;
 RX MEDLINE=94132007; PubMed=8300574;
 RA Nagashima K.V., Matsuura K., Ohnaya S., Shimada K.;
 RT "Primary structure and transcription of genes encoding B870 and
 RT photosynthetic reaction center apoproteins from Rubrivivax
 RT gelatinosus.";
 RL J. Biol. Chem. 269:2477-2484(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL144;
 RA Igarashi N., Shimada K., Matsuura K., Nagashima K.V.;
 RT "Photosynthetic gene cluster in purple bacterium, Rubrivivax
 RT gelatinosus.";

RL (In) Garab G. (eds.);
 RL Photosynthesis;
 RL mechanisms and effects (Proceedings of the 11th international congress
 RL on photosynthesis), pp.4:2889-2892, Kluwer Academic Publishers,
 RL Dordrecht (1999).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL144;
 RX MEDLINE=20031519; PubMed=10563807;
 RA Menin L., Yoshida M., Jaquinod M., Nagashima K.V., Matsuura K.,
 RA Parot P., Vermeiglio A.;
 RT "Dark aerobic growth conditions induce the synthesis of a high
 RT midpoint potential cytochrome c8 in the photosynthetic bacterium
 RT Rubrivivax gelatinosus.";
 RL Biochemistry 38:15238-15244(1999).
 DR EMBL; ABC34704; BAA94057.1; -;
 DR PIR; T50904; T50904.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0009058; P:biosynthesis; IEA.
 DR InterPro; IPR003672; Coby/Mg_chltase.
 DR Pfam; PF02514; Coby-Mg_chel; 1.
 KW Transference.
 SQ SEQUENCE 1236 AA; 134729 MW; 84051C045638520C CRC64;

Query Match 32.9%; Score 55; DB 2; Length 1236;
 Best Local Similarity 34.4%; Pred. No. 45;
 Matches 11; Conservative 6; Mismatches 9; Indels 6; Gaps 1;

OY 3 AHQSMETPTDNPINPAWAG-----RGIRPV 28
 :|||:|
 DB 1112 SEQVALETRMLNPKWYEGMLHGYGVQR 1143
 :|||:|

RESULT 12

ID O9KRY1 PRELIMINARY; PRT; 503 AA.
 AC O9KRY1;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Sun/nucleolar protein family protein.
 GN VCI502.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=566;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=El Tor N16961 / Serotype O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae.";
 RL Nature 406:477-483(2000).
 DR ENBL; AE004228; AAP94657.1; -;
 DR PIR; A82193; A82193.
 DR TIGR; VCI502; -;
 DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. .; IEA.
 DR InterPro; IPR000051; SAM_bind.
 DR InterPro; IPR001678; Sun_Nop1/Nop2.
 DR Pfam; PF01189; Noll_Nop2_Sun; 1.
 DR TIGRFAMs; TIGR00446; nop2p; 1.
 DR PROSITE; PS01153; NOLL_NOP2_SUN; 1.
 KW Complete proteome.
 SQ SEQUENCE 503 AA; 55997 MW; 2ABD94A235C9E48 CRC64;

Query Match 32.6%; Score 54.5; DB 16; Length 503;

Best Local Similarity 36.7%; Pred. No. 20;
Matches 11; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

QY 1 SRAHQHSMETPDPINPAWYAGRGIRPVGR 30
DB 435 SSSASHVELDTQAR-EWFGDRDVRPEGQ 463

RESULT 13
Q7UPG7 PRELIMINARY; PRT; 1120 AA.
AC Q7UPG7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN RB6951.
OS Rhodopirellula baltica.
OC Bacteria; planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294145; CAD75095.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1120 AA; 123221 MW; 67C757B73421BF9B CRC64;

Query Match 32.3%; Score 54; DB 16; Length 1120;
Best Local Similarity 64.3%; Pred. No. 56;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 14 DINPAWYAGRGIRP 27
DB 202 EISPAWAKXGIRP 215

RESULT 14
Q08689 PRELIMINARY; PRT; 176 AA.
AC Q08689;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Chromosome XV reading frame ORF YOR253W.
GN NAT5 OR ARD2 OR YOR253W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
CX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA MTPS;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97298311; PubMed=9153759;
RA Jauniaux J.C., Poiray R.;
RT "Sequencing analysis of a 36.8 kb fragment of yeast chromosome XV
reveals 26 open reading frames including SEC63, CDC31, SUG2, GCD1,
REL2, FNT1, PAC1 and VPH1.";
RL Yeast 13:483-487(1997).
DR EMBL; Z75161; CAA99475.1; -.
DR PIR; S67150; S67150.
DR SGD; S0005779; NAT5.
DR GO; GO:0005737; Cytoplasm; IC.

DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR000182; GCN5acetyl_trans.
DR Pfam; PF00583; Acetyltransf. 1.
DR PROSITE; PS00190; CYTOCHROME C; 1.
SQ SEQUENCE 176 AA; 19727 MW; 4F09DC597A690BA0 CRC64;

Query Match 32.0%; Score 53.5; DB 3; Length 176;
Best Local Similarity 31.2%; Pred. No. 9.3;
Matches 10; Conservative 5; Mismatches 14; Indels 3; Gaps 1;

QY 1 SRAHQHSMETPDPINPAWYAGRGIRPVG 29
DB 120 SECHQNVFVYLPVAVDDLTQKQWIAHGFEQVG 151

RESULT 15
Q9RYP2 PRELIMINARY; PRT; 376 AA.
AC Q9RYP2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Adenine deaminase-related protein.
GN DRA0268.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
Moffat K.S., Qin H., Jiang L., Ramphile W., Crosby M., Shen M.,
Vamathevan J.J., Lam P., McDonald L., Utterback L., Zaleski C.,
Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AE001863; AAF12376.1; -.
DR PIR; C75580; C75580.
DR TIGR; DRA0268; -.
KW Complete proteome.
SQ SEQUENCE 376 AA; 39845 MW; 7AB7FF32F8C45651 CRC64;

Query Match 31.7%; Score 53; DB 16; Length 376;
Best Local Similarity 40.0%; Pred. No. 25;
Matches 12; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 1 SRAHQHSMETPDPINPAWYAGRGIRPVGR 30
DB 177 ARALAHAPDVSRPDRHRACGQRRRPAQR 206

Search completed: August 12, 2004, 14:48:58
Job time : 44.7151 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:36 ; Search time 13.6977 Seconds
(without alignments)
217.697 Million cell updates/sec

Title: US-09-700-643A-1
Perfect score: 167
Sequence: 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGRX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*

1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	157	94.0	83	JC7607	prolactin-releasing
2	56.5	33.8	303	AH2016	hypothetical protein
3	55	32.9	1236	T50504	Mg protoporphyrin
4	54.5	32.6	503	A82193	sun/molecular prot
5	53.5	32.0	176	S67150	hypothetical protein
6	53	31.7	376	C75580	adenine deaminase
7	52	31.1	73	A98137	hypothetical protein
8	52	31.1	308	AH2896	transcription regu
9	52	31.1	308	B97672	fok family protein
10	52	31.1	419	AH3166	hypothetical protein
11	52	31.1	1232	T31462	probable magnesium
12	51.5	30.8	664	F83376	conserved hypother
13	51	30.5	294	T21075	hypothetical protein
14	51	30.5	430	B69009	conserved hypother
15	51	30.5	798	S11210	probable untr prote
16	50	29.9	527	T33175	hypothetical protein
17	49.5	29.6	333	AB2664	conserved hypother
18	49.5	29.6	401	A97446	hypothetical protein
19	49.5	29.6	790	T47959	hypothetical protein
20	49	29.3	128	S76955	hypothetical protein
21	48.5	29.0	254	S76914	hypothetical protein
22	48.5	29.0	548	T47548	hypothetical protein
23	48	28.7	72	E91002	probable regulator
24	48	28.7	220	C83292	probable glutathio
25	48	28.7	314	B70569	hypothetical protein
26	48	28.7	348	T21648	hypothetical protein
27	48	28.7	365	T39098	hypothetical protein
28	48	28.7	424	B38176	samb protein - Sal
29	48	28.7	637	D87559	sensory box histid

30	48	28.7	772	2	T07958	protoporphyrin IX
31	48	28.7	1328	2	AE2351	protoporphyrin IX
32	48	28.7	1331	2	S75000	protoporphyrin IX
33	48	28.7	1379	2	S37310	protoporphyrin IX
34	48	28.7	1380	2	S64721	protoporphyrin IX
35	48	28.7	1381	2	S71288	protoporphyrin IX
36	48	28.7	1382	2	T01789	protoporphyrin IX
37	48	28.7	1383	2	T07126	magnesium chelatase
38	47.5	28.4	777	2	AD0962	biotin sulfoxide r
39	47	28.1	406	2	I46535	11-beta-hydroxyste
40	47	28.1	418	2	JC7588	exo-alpha-sialidas
41	47	28.1	455	2	D70885	probable aldc prot
42	47	28.1	785	2	F69099	sensory transducti
43	47	28.1	1084	2	T33759	hypothetical prote
44	46.5	27.8	345	2	D84012	N-acetylglutamate
45	46.5	27.8	479	2	AB0729	conserved hypother

ALIGNMENTS

RESULT 1

JC7607

prolactin-releasing peptide - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001

C:Accession: JC7607

R:Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Hosoya, T.

Biogem. Biophys. Res. Commun. 281, 53-56, 2001

A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene: MuJ

A:Reference number: JC7607; MUID:21092785; PMID:11178959

A:Contents: Spleen

A:Accession: JC7607

A:Molecule type: DNA

A:Residues: 1-83 <YAM>

A:Cross-references: DBJ:AB040612; DBJ:AB040613

C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior p

release, and stimulation of ACTH secretion from the pituitary.

C:Genetics:

A:Gene: PrRP

A:Introns: 33/1

Query Match 94.0%; Score 157; DB 2; Length 83;
Best Local Similarity 93.3%; Pred. No. 6e-16;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGR 30

DB 22 SRAHQSHMEIRTPDINPAWYAGRGIRPVGR 51

RESULT 2

AH2016

hypothetical protein all1686 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 24-Nov-2003

C:Accession: AH2016

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi

Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AH2016

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-303 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA078052.1; PID:gl71355506; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all1696

C:Superfamily: tartrate-resistant acid phosphatase

Query Match 33.8%; Score 56.5; DB 2; Length 303;
 Best Local Similarity 51.6%; Pred. No. 1.5;
 Matches 16; Conservative 2; Mismatches 5; Gaps 2;

Qy 4 HQHSMETPDINPANY-----AGRGIRPVGR 30
 Db 226 HEHSYE-RTAIDGTTTLTCGAGAGNRPVGR 255

RESULT 3
 T50904
 Mg protoporphyrin methyl transferase [imported] - Rubrivivax gelatinosus
 C:Species: Rubrivivax gelatinosus
 C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 02-Sep-2000
 C:Accession: T50904
 R:Nagashima, K.V.; Igarashi, N.; Harada, J.; Nagashima, S.; Matsuura, K.; Shimada, K.
 submitted to the EMBL Data Library, November 1999
 A:Description: Determination of Nucleotide Sequences of Rubrivivax gelatinosus Photosynthetic
 A:Reference number: 225270
 A:Accession: T50904
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-1236 <NAG>
 A:Cross-references: EMBL:AB034704; PIDN:BAA94057.1
 A:Experimental source: strain IL144
 C:Genetics:
 A:Gene: bchH
 C:Superfamily: Rhodobacter capsulatus magnesium-protoporphyrin O-methyltransferase

Query Match 32.9%; Score 55; DB 2; Length 1236;
 Best Local Similarity 34.4%; Pred. No. 11;
 Matches 11; Conservative 6; Mismatches 9; Indels 6; Gaps 1;

Qy 3 AHQHSMEIRTPDINPANYAG-----RGIRPV 28
 Db 1112 SEQVALETATMLNPKWYEGMLEHGYGVQI 1143

RESULT 4
 A82193
 Sun/nucleolar protein family protein VC1502 [imported] - Vibrio cholerae (strain N16961)
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C:Accession: A82193
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MUID:20406833; PMID:10952301
 A:Accession: A82193
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-503 <HEI>
 A:Cross-references: GB:AE004228; GB:AE003852; NID:G9655997; PIDN:AAF94657.1; GSPDB:GN001
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VC1502
 A:Map position: 1

Query Match 32.6%; Score 54.5; DB 2; Length 503;
 Best Local Similarity 36.7%; Pred. No. 5;
 Matches 11; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

Qy 1 SRAHQHSMETPDINPANYAGRGIRPVGR 30
 Db 435 SSSASHSELDTTQAR-EWFMGRDVRPEQG 463

RESULT 5
 S67150
 hypothetical protein YOR253w - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein O5315

C:Species: Saccharomyces cerevisiae
 C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 19-Apr-2002
 C:Accession: S67150
 R:Jauniaux, J.C.; Poirey, R.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S67143
 A:Accession: S67150
 A:Molecule type: DNA
 A:Residues: 1-176 <JAU>
 A:Cross-references: EMBL:Z75161; NID:gl420572; PID:e252411; PID:gl420573; GSPDB:GN00015;
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: MIP5:YOR253w
 A:Cross-references: SGD:S0005779
 A:Map position: 15R

Query Match 32.0%; Score 53.5; DB 2; Length 176;
 Best Local Similarity 31.2%; Pred. No. 2.3;
 Matches 10; Conservative 5; Mismatches 3; Gaps 1;

Qy 1 SRAHQHSMETPDINPANYAGRGIRPVGR 29
 Db 120 SECHQHNVFYLPVAVDDLTKQWFIAGFEQVG 151

RESULT 6
 C75580
 adenine deaminase-related protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C:Accession: C75580
 R:White, C.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: C75580
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-376 <WHI>
 A:Cross-references: GB:AE001863; GB:AE001825; NID:G6460670; PIDN:AAF12376.1; PID:G646067
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DRA0268
 A:Map position: 2

Query Match 31.7%; Score 53; DB 2; Length 376;
 Best Local Similarity 40.0%; Pred. No. 6.1;
 Matches 12; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

Qy 1 SRAHQHSMETPDINPANYAGRGIRPVGR 30
 Db 177 ARALAHAPVSDRDRHRACGAGORRRPAGR 206

RESULT 7
 A98137
 hypothetical protein AGR_L109 [imported] - Agrobacterium tumefaciens (strain C58, Cereon
 C:Species: Agrobacterium tumefaciens
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
 C:Accession: A98137
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, P.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
 A:Reference number: A97359; MUID:21608551; PMID:11743194
 A:Accession: A98137
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-73 <KUR>
 A:Cross-references: GB:AE007870; PIDN:AAK88619.1; PID:gl5158338; GSPDB:GN00170
 C:Genetics:

R;Stover, C.K.; Pham, X.Q.; Erwin, A.

.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: F83376
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-664 <STO>
A;Cross-references: GB:AE004642; GB:AE004091; NID:G9948163; PIDN:AAG05539.1; GSPDB:GN001
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA2151

Query Match 30.8%; Score 51.5; DB 2; Length 664;
Best Local Similarity 55.8%; Pred. No. 19;
Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 10 IRTPDINPAWYAGRGIRP 27
: ||||| :
Db 476 VNTPDINP-WFLQSGRP 492
: ||||| :
: ||||| :
: ||||| :

RESULT 13
T21075
hypothetical protein F19H6.1 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C;Accession: T21075; T2124
R;McMurray, A.
submitted to the EMBL Data Library, August 1995
A;Reference number: Z19368
A;Accession: T21075
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-294 <WIL>
A;Cross-references: EMBL:Z50873; PIDN:CAA90762.1; GSPDB:GN000028; CESP:F19H6.1
A;Experimental source: clone F17E5
R;McMurray, A.
submitted to the EMBL Data Library, November 1995
A;Reference number: Z19378
A;Accession: T21224
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-294 <W12>
A;Cross-references: EMBL:Z68115; PIDN:CAA92169.1; GSPDB:GN000028; CESP:F19H6.1
A;Experimental source: clone F19H6
C;Genetics:
A;Gene: CESP:F19H6.1
A;Map position: X
A;Introns: 55/3; 85/2; 113/3; 150/1; 192/3; 266/2
C;Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 30.5%; Score 51; DB 2; Length 294;
Best Local Similarity 44.8%; Pred. No. 9.2;
Matches 13; Conservative 2; Mismatches 12; Indels 2; Gaps 1;

QY 3 AHQSHMEIRTPDINPA--WYAGRGIRPVG 29
: ||||| :
Db 139 AHMHSKRIMRDIKPANVFITGNGIVKLG 167
: ||||| :
: ||||| :
: ||||| :

RESULT 14
B69009
conserved hypothetical protein MTH1070 - *Methanobacterium thermoautotrophicum* (strain De
C;Species: *Methanobacterium thermoautotrophicum*
C;Date: 29-Jan-1999 #sequence_revision 29-Jan-1999 #text_change 23-Jul-1999
C;Accession: B69009
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of *Methanobacterium thermoautotrophicum* Delta H: funct
A;Reference number: A69000; MUID:98037514; PMID:9371463

A;Accession: B69009
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-430 <MTH>
A;Cross-references: GB:AE000877; GB:AE000666; NID:G2622157; PIDN:AAB85559.1; PID:G262216
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH1070
C;Superfamily: conserved hypothetical protein MTH1070

Query Match 30.5%; Score 51; DB 1; Length 430;
Best Local Similarity 69.2%; Pred. No. 14;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 14 DINPAWYAGRGIR 26
: ||||| :
Db 191 DINPEWVAGRACR 203
: ||||| :
: ||||| :
: ||||| :

RESULT 15
S11210
probable unr protein - rat
C;Species: *Rattus norvegicus* (Norway rat)
C;Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999
C;Accession: S11210
R;Jeffers, M.; Paciucci, R.; Fellicer, A.
Nucleic Acids Res. 18, 4891-4899, 1990
A;Title: Characterization of unr; a gene closely linked to N-ras.
A;Reference number: S11210; MUID:90370473; PMID:2204029
A;Accession: S11210
A;Molecule type: mRNA
A;Residues: 1-798 <JEF>
A;Cross-references: EMBL:X52311; NID:G57454; PIDN:CAA36549.1; PID:G57455
C;Keywords: DNA binding

Query Match 30.5%; Score 51; DB 2; Length 798;
Best Local Similarity 43.5%; Pred. No. 27;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 6 HSMETIRTPDINPAWYAGRGIRPV 28
: ||||| :
Db 583 HSYNGITEANPTYSKVRPL 605
: ||||| :
: ||||| :
: ||||| :

Search completed: August 12, 2004, 14:50:24
Job time : 14.6977 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 8.11047 Seconds
(without alignments)
199.024 Million cell updates/sec

Title: US-09-700-643A-1

Perfect score: 167

Sequence: 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGRX 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	166	99.4	98	1 PRRP_BOVIN	P81264 bos taurus
2	157	94.0	83	1 PRRP_RAT	P81278 rattus norv
3	152	91.0	87	1 PRRP_HUMAN	P81277 homo sapien
4	61	36.5	428	1 NER3_BOVIN	Q97859 bos taurus
5	56	33.5	428	1 NER3_HUMAN	Q9u49 homo sapien
6	52	31.1	413	1 EX7L_COREF	Q8f6p1 corynebacte
7	52	31.1	417	1 EX7L_CORGL	Q8nrm3 corynebacte
8	51	30.5	798	1 UNR_RAT	P18395 rattus norv
9	50	29.9	288	1 Y587_PASMU	Q9cns56 pasteurella
10	48.5	29.0	239	1 6PGL_SYNY3	P74618 synchocyst
11	48	28.7	434	1 SAME_SALTY	P23832 salmonella
12	47.5	28.4	1083	1 T2D3_HUMAN	O00268 homo sapien
13	47	28.1	316	1 XERC_VIBVU	Q7za19 vibrio vuln
14	47	28.1	406	1 DHI2_RABIT	P51976 oryctolagus
15	47	28.1	418	1 NER3_RAT	Q99pw5 rattus norv
16	47	28.1	442	1 TRB3_METMA	Q8q001 methanosarc
17	47	28.1	1308	1 M4K6_MOUSE	Q9jms2 mus musculu
18	46.5	27.8	345	1 ARGC_BACHD	Q9k8v2 bacillus ha
19	46.5	27.8	770	1 AVPS_ARATH	P31414 arabidopsis
20	46	27.5	277	1 DAPF_CORGL	Q8nfp73 corynebacte
21	46	27.5	342	1 Y762_METJA	Q58172 methanococ
22	46	27.5	347	1 Y576_METJA	Q57996 methanococ
23	46	27.5	402	1 EX7L_STRCO	Q9f6m3 streptomyce
24	46	27.5	510	1 YCGE_ECOLI	P29013 escherichia
25	46	27.5	798	1 UNR_HUMAN	Q75534 homo sapien
26	46	27.5	836	1 TGM1_RABIT	P22758 oryctolagus
27	45.5	27.2	272	1 TRA2_DROVI	O02008 drosophila
28	45.5	27.2	374	1 YHHJ_ECOLI	P31993 escherichia
29	45.5	27.2	411	1 PUR7_ARATH	P38025 arabidopsis
30	45.5	27.2	670	1 SUV1_ARATH	Q9ff80 arabidopsis
31	45.5	27.2	704	1 LIPM_NEIMB	Q05013 neisseria m
32	45	26.9	387	1 HISS_HAINI	Q9hg50 halobacteri
33	45	26.9	418	1 NER3_MOUSE	Q9jnm7 mus musculu

34	45	26.9	482	1	R167_YEAST
35	45	26.9	817	1	TGM1_HUMAN
36	45	26.9	824	1	TGM1_RAT
37	45	26.9	962	1	UVRA_METH
38	45	26.9	1033	1	ACAC_ARATH
39	44.5	26.6	407	1	Y116_MYCTU
40	44.5	26.6	474	1	ATZA_PSESD
41	44.5	26.6	772	1	LMBT_HUMAN
42	44.5	26.6	777	1	BISC_ECOLI
43	44.5	26.6	895	1	RA50_THEVO
44	44	26.3	246	1	NPD_DEIRA
45	44	26.3	358	1	PLCA_ERWCH

ALIGNMENTS

RESULT 1
PRRP_BOVIN STANDARD; PRT; 98 AA.
ID PRRP_BOVIN
AC P81264;
DT/ 30-MAY-2000 (Rel. 39, Created)
DT/ 30-MAY-2000 (Rel. 39, Last sequence update)
DT/ 28-FEB-2003 (Rel. 41, Last annotation update)
DE Prolactin-releasing peptide precursor (PRP) (Prolactin-releasing hormone) (Contains: Prolactin-releasing peptide PrRP31; Prolactin-releasing peptide PrRP20).
GN PRH.

OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.
RC TISSUE=Brain;
RX MEDLINE=98268781; PubMed=9607765;
RA Hinuma S., Habata Y., Fujii R., Kawanata Y., Hosoya M., Fukusumi S., Kitada K., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.,
RA "A prolactin-releasing peptide in the brain."
RL Nature 393:272-276(1998)
CC -!- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.
CC -!- TISSUE SPECIFICITY: Medulla oblongata and hypothalamus.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).

CC EMBL; AB015417; BAA29025.1;
KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.
FT SIGNAL 1 22
FT PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRRP31.
FT PEPTIDE 33 53 PROLACTIN-RELEASING PEPTIDE PRRP20.
FT PROPEP 58 98
FT MOD RES 53 53
SQ SEQUENCE 98 AA; 10544 MW; 0BAC35A13B0FA908 CRC64;
AMIDATION (G-54 PROVIDE AMIDE GROUP).

Query Match 99.4%; Score 166; DB 1; Length 98;
Best Local Similarity 100.0%; Pred. NO. 1.7e-17;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGR 30

DB 23 SRAHQSHMEIRTPDINPAWYAGRGIRPVGR 52

RESULT 2	PRRP RAT	STANDARD;	PRT;	83 AA.
ID	PRRP RAT	STANDARD;	PRT;	83 AA.
AC	P81278; Q8K3Y0;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Prolactin-releasing peptide precursor (PRP) (Prolactin-releasing hormone) [Contains: Prolactin-releasing peptide PRRP31; Prolactin-releasing peptide PRP20].			
GN	PRH.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OC	NCBI_TaxId:10116;			
OC	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RC	TISSUE=Brain;			
RC	MEDLINE=98268781; PubMed=9607765;			
RX	Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,			
RA	Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,			
RA	Kurokawa T., Nishimura O., Onda H., Fujino M.;			
RT	"A prolactin-releasing peptide in the brain.";			
RT	Nature 393:272-276(1998).			
RN	[2]			
RN	SEQUENCE FROM N.A. (ISOFORM 2).			
RC	STRAIN=Sprague-Dawley; TISSUE=Hypothalamus;			
RC	Anderson S.T., Kokay I.C., Lang T., Grattan D.R., Curlewis J.D.;			
RT	"Quantitation of prolactin-releasing peptide (PRP) mRNA expression in			
RT	specific brain regions during the rat oestrous cycle and in			
RT	lactation.";			
RT	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RN	TISSUE SPECIFICITY.			
RP	MEDLINE=9942652; PubMed=10498338;			
RX	Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,			
RA	Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,			
RA	Sumino Y., Fujino M.;			
RT	"Tissue distribution of prolactin-releasing peptide (PrRP) and its			
RT	receptor.";			
RT	Regul. Pept. 83:1-10(1999).			
CC	!- FUNCTION: Stimulates prolactin (PRL) release and regulates the			
CC	expression of prolactin through its receptor GPR10. May stimulate			
CC	lactotrophs directly to secrete PRL.			
CC	!- ALTERNATIVE PRODUCTS:			
CC	Event=alternative splicing; Named isoforms=2;			
CC	Name=1;			
CC	ISOID=P81278-1; Sequence=Displayed;			
CC	Name=2;			
CC	ISOID=P81278-2; Sequence=VSP_004370;			
CC	TISSUE SPECIFICITY: Widely expressed, with highest levels in			
CC	medulla oblongata and hypothalamus.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; AB015418; BAA29026.1; -			
DR	EMBL; AF521930; AAM82154.1; -			
DR	PIR; JC7607; JC7607.			
DR	Hormone; Amidation; Signal; Cleavage on pair of basic residues;			
KW	Alternative splicing			
FT	SIGNAL 1 21			BY SIMILARITY.
FT	PEPTIDE 22 52			PROLACTIN-RELEASING PEPTIDE PRRP31.
FT	PEPTIDE 33 52			PROLACTIN-RELEASING PEPTIDE PRRP20.
FT	PROPEP 57 83			
FT	MOD_RES 52 52			
FT	VARSPPLIC 33 83			AMIDATION (G-53 PROVIDE AMIDE GROUP). TPDINPAWTKGRRVFRGRRRATPRDVTGLGSLCPL DGTGVSORG -> SCIVNGKPISTSPHPTSSOMPP (in

```

isoform 2).
/FTIG=VSPF004370.
DOC75A264EEB4F29 CRC64;

SEQUENCE 83 AA; 9215 MW; 94.0%; Score 157; DB 1; Length 83;
Best Local Similarity 93.3%; Pred. No. 3e-16;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Query Match
1 SRAQHSMEIRTPDINPAWYAGRGIRPVGR 30
|||||
22 SRAQHSMETRTPDINPAWYTGIRPVGR 51
|||||

RESULT 3
PRRP HUMAN STANDARD; PRT; 87 AA.
ID PRRP HUMAN
AC P81277;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing
DE hormone) [Contains: Prolactin-releasing peptide PRP31; Prolactin-
DE releasing peptide PRP20].
GN PRH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RX [1]
RN "A prolactin-releasing peptide in the brain."
RT Nature 393:272-276(1998).
RL [2]
TISSUE SPECIFICITY.
RP MEDLINE=99426652; PubMed=10498338;
RX Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,
RX Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,
RX Sumano Y., Fujino M.;
RA "Tissue distribution of prolactin-releasing peptide (PrRP) and its
RX receptor."
RT Regul. Pept. 83:1-10(1999).
CC -!- FUNCTION: Stimulates prolactin (PRL) release and regulates the
CC expression of prolactin through its receptor GPR10. May stimulate
CC lactotrophs directly to secrete PRL.
CC -!- TISSUE SPECIFICITY: Medulla oblongata and hypothalamus.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AB015419; BAA29027.1; -.
DR DR
DR MIM; 602663; -.
DR GO; GO:0005180; F:peptide hormone; TAS.
DR Hormone; Amidation; Signal.
KW SIGNAL 1 22
FT BY SIMILARITY.
FT PEPTIDE 23 53
FT PEPTIDE 34 53
FT PROPEP 58 87
FT MOD_RES 53 53
FT SEQUENCE 87 AA; 9639 MW; 229A2F3P50CF981B CRC64;

Query Match
152; DB 1; Length 87;
Best Local Similarity 90.0%; Pred. No. 1.7e-15;
Matches 27; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```



```
OY 1 SRAHQHMEIRTPDINPAWYAGRGIRPVGR 30
DB 23 SRTSHSMEIRTPDINPAWYASRGIRPVGR 52

RESULT 4
NER3 BOVIN
ID NER3 BOVIN STANDARD; PRT; 428 AA.
AC O97859;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sialidase 3 (EC 3.2.1.18) (Membrane sialidase) (Ganglioside sialidase)
DE (N-acetyl-alpha-neuraminidase 3).
GN NEU3.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN NCBI_TaxID=9913;
[1]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.
RC TISSUE=Brain;
RX MEDLINE=99143165; PubMed=9989745;
RA Miyagi T., Wada T., Iwamatsu A., Hata K., Yoshikawa Y., Tokuyama S.,
RA Sawada M.;
RT "Molecular cloning and characterization of a plasma membrane-
associated sialidase specific for gangliosides.";
RL J. Biol. Chem. 274:5004-5011(1999).
CC -!- FUNCTION: Plays a role in modulating the ganglioside content of
the lipid bilayer at the level of membrane-bound sialyl
glycoconjugates.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2-3)-, alpha-(2-6)-,
alpha-(2-8)-glycosidic linkages of terminal sialic residues in
oligosaccharides, glycoproteins, glycolipids, colominic acid and
synthetic substrates.
CC -!- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed in brain.
CC -!- SIMILARITY: Belongs to family 33 of glycosyl hydrolases.
CC -!- SIMILARITY: Contains 3 BNR repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
EMBL: AB008184; BAA75071.1; .
DR InterPro: IPR002860; GH_BNR.
DR Pfam: PF02012; BNR; 3.
KW Hydrolase; Glycosidase; Membrane; Repeat.
FT REPEAT 129 140 BNR 1.
FT REPEAT 203 214 BNR 2.
FT REPEAT 254 265 BNR 3.
FT SITE 24 27 FRIP MOTIF.
FT ACT_SITE 25 25 By similarity.
FT ACT_SITE 45 45 Potential.
FT ACT_SITE 50 50 Potential.
FT ACT_SITE 87 87 Potential.
FT ACT_SITE 225 225 Potential.
FT ACT_SITE 245 245 Potential.
FT ACT_SITE 341 341 By similarity.
FT ACT_SITE 371 371 Potential.
FT ACT_SITE 388 388 Potential.
SQ SEQUENCE 428 AA; 47916 MW; 418B34E3245A8F21 CRC64;

Query Match 36.5%; Score 61; DB 1; Length 428;
Best Local Similarity 40.7%; Pred. NO. 0.21;
Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

OY 2 RARQHSMEIRTPDINPAWYAGRGIRPV 28
DB 195 RARPHSLMYSDDLGLGATWHGRLIKPM 221

RESULT 5
NER3 HUMAN
ID NER3 HUMAN STANDARD; PRT; 428 AA.
AC Q9UQ49; Q9NQE1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sialidase 3 (EC 3.2.1.18) (Membrane sialidase) (Ganglioside sialidase)
DE (N-acetyl-alpha-neuraminidase 3).
GN NEU3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99335353; PubMed=10405317;
RA Wada T., Yoshikawa Y., Tokuyama S., Kuwabara M., Akita H., Miyagi T.;
RA Croci G., Preti A., Ballabio A., Tettamanti G., Borsani G.;
RT "Cloning, expression, and chromosomal mapping of a human ganglioside
sialidase.";
RL Biochem. Biophys. Res. Commun. 261:21-27(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX PubMed=10861246;
RA Monti E., Bassi M.T., Papini N., Riboni M., Manzoni M., Venerando B.,
RA Croci G., Preti A., Ballabio A., Tettamanti G., Borsani G.;
RT "Identification and expression of NEU3, a novel human sialidase
associated to the plasma membrane.";
RL Biochem. J. 349:343-351(2000).
CC -!- FUNCTION: Plays a role in modulating the ganglioside content of
the lipid bilayer at the level of membrane-bound sialyl
glycoconjugates.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2-3)-, alpha-(2-6)-,
alpha-(2-8)-glycosidic linkages of terminal sialic residues in
oligosaccharides, glycoproteins, glycolipids, colominic acid and
synthetic substrates.
CC -!- SUBCELLULAR LOCATION: Membrane-associated.
CC -!- TISSUE SPECIFICITY: Highly expressed in skeletal muscle, testis,
adrenal gland and thymus, followed by pancreas, liver, heart and
thymus. Weakly expressed in kidney, placenta, brain and lung.
CC -!- MISCELLANEOUS: Optimum pH is 3.8.
CC -!- SIMILARITY: Belongs to family 33 of glycosyl hydrolases.
CC -!- SIMILARITY: Contains 3 BNR repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
EMBL: AB008185; BAA82611.1; .
DR EMBL: Y18563; CAB96131.1; ALT_INIT.
DR Genew: HGNC:7760; NEU3.
DR MIM: 604617; .
DR GO: GO:0005887; C: integral to plasma membrane; TAS.
DR GO: GO:0003824; F: catalytic activity; TAS.
DR GO: GO:0006689; P: ganglioside catabolism; TAS.
DR InterPro: IPR002860; GH_BNR.
DR Pfam: PF02012; BNR; 3.
KW Hydrolase; Glycosidase; Membrane; Repeat.
FT REPEAT 129 140 BNR 1.
FT REPEAT 203 214 BNR 2.
FT REPEAT 254 265 BNR 3.
FT SITE 24 27 FRIP MOTIF.
```

```

FT ACT_SITE 25 25 BY similarity.
FT ACT_SITE 45 45 Potential.
FT ACT_SITE 50 50 Potential.
FT ACT_SITE 87 87 Potential.
FT ACT_SITE 225 225 Potential.
FT ACT_SITE 245 245 Potential.
FT ACT_SITE 340 340 By similarity.
FT ACT_SITE 370 370 Potential.
FT ACT_SITE 387 387 Potential.
SQ SEQUENCE 428 AA; 48252 MW; 35D1DD9359A78C98 CRC64;

Query Match 33.5%; Score 56; DB 1; Length 428;
Best Local Similarity 37.0%; Pred. No. 1.2;
Matches 10; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 2 RAHQSHMEITPTINAWAGRGIRPV 28
DB 195 KTRPHSLMTYDDLGVTHHGRILRPM 221

RESULT 6
EX7L_COREF STANDARD; PRT; 413 AA.
AC Q8FQPL;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
DE (Exonuclease VII large subunit).
DE XSEA OR CGL078.
GN Corynebacterium efficiens.
OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y8-314 / AJ 12310 / DSM 44549 / JCM 11189;
RX MEDLINE=22723752; Pubmed=12840036;
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gofobori T.;
RT Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens."
RL Genome Res. 13:1572-1579(2003).
CC -!- FUNCTION: Bidirectionally degrades single-stranded DNA into large
CC acid-insoluble oligonucleotides, which are then degraded further
CC into small acid-soluble oligonucleotides (By similarity).
CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.
CC -!- SUBUNIT: Heterooligomer composed of large and small subunits (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the xseA family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP005217; BAC17888.1; ALT_INIT.
CC HAMAP; MF_00378; -; 1.
CC InterPro; IPR003753; Exonuc_VII_L.
CC InterPro; IPR008994; Nucleic_acid_OB.
CC Pfam; PF02601; Exonuc_VII_L; 1.
CC TIGRPFAMs; TIGR00237; xseA; 1.
CC HydroLase; Nuclease; Exonuclease; Complete proteome.
CC SEQUENCE 413 AA; 45171 MW; D3BAD9687C6E85D CRC64;

Query Match 31.1%; Score 52; DB 1; Length 413;
Best Local Similarity 57.1%; Pred. No. 4.4;
Matches 12; Conservative 1; Mismatches 0; Indels 8; Gaps 1;

QY 17 PAWYAGRG-----IRPVG 29
DB 94 PAFYAGRGFSLWVTDIRPVG 110

RESULT 7
EX7L_COREF STANDARD; PRT; 417 AA.
AC Q8NRM3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
DE (Exonuclease VII large subunit).
DE XSEA OR CGL1025.
GN Corynebacterium glutamicum (Brevibacterium flavum).
OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RL "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Bidirectionally degrades single-stranded DNA into large
CC acid-insoluble oligonucleotides, which are then degraded further
CC into small acid-soluble oligonucleotides (By similarity).
CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.
CC -!- SUBUNIT: Heterooligomer composed of large and small subunits (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the xseA family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP005277; BAB98418.1; -;
CC HAMAP; MF_00378; -; 1.
CC InterPro; IPR003753; Exonuc_VII_L.
CC InterPro; IPR008994; Nucleic_acid_OB.
CC Pfam; PF02601; Exonuc_VII_L; 1.
CC TIGRPFAMs; TIGR00237; xseA; 1.
CC HydroLase; Nuclease; Exonuclease; Complete proteome.
CC SEQUENCE 417 AA; 45582 MW; B32CD9286C173C34 CRC64;

Query Match 31.1%; Score 52; DB 1; Length 417;
Best Local Similarity 57.1%; Pred. No. 4.4;
Matches 12; Conservative 1; Mismatches 0; Indels 8; Gaps 1;

QY 17 PAWYAGRG-----IRPVG 29
DB 94 PAFYAGRGFSLWVTDIRPVG 114

RESULT 8
UNR_RAT STANDARD; PRT; 798 AA.
ID UNR_RAT
AC P18395;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE UNR protein.
GN UNR.

```



```
DR EMBL; D90916; BAA18726.1; ALT INIT.
DR InterPro; IPR006148; Gluc gal isom.
DR InterPro; IPR005900; Phosphoglucanlac.
DR Pfam; PF01182; Glucoamine iso; 1.
DR TIGRfam; TIGR01198; pgl; 1.
KW Hydrolase; Complete proteome.
FT INIT_MET 0
SQ SEQUENCE 239 AA; 26351 MW; 9C64A0A342325917 CRC64;

Query Match 29.0%; Score 48.5; DB 1; Length 239;
Best Local Similarity 42.3%; Pred. No. 8.1;
Matches 11; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 5 QHSM-BIRTPDNPANVYAGRGIRPVG 29
Db 198 QHAGLGFAPADPQYPARFIQPOG 223

RESULT 11
SAME_SALT SAMB_SALT STANDARD; PRT; 424 AA.
AC P23832;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Smb protein.
GN Smb.
OS Salmonella typhimurium.
OG Plasmid 60-mDa cryptic.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=91123176; PubMed=1991707;
RA Nohmi T., Hakura A., Nakai Y., Watanabe M., Murayama S.Y.,
RA Sofuni T.;
RT "Salmonella typhimurium has two homologous but different umuC
RT operons: cloning of a new umuC-like operon (samb) present in a
RT 60-megadalton cryptic plasmid of S. typhimurium.";
RL J. Bacteriol. 173:1051-1063(1991).
CC -!- FUNCTION: Involved in UV protection and mutation.
CC -!- SIMILARITY: Belongs to the DNA polymerase type-Y family.
CC -!- SIMILARITY: Contains 1 umuC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; D90202; BAA14226.1; --
DR PIR; B38176; B38176.
DR HAMAP; MF 01113; atypical; 1.
DR InterPro; IPR001126; UMUC_like.
DR Pfam; PF00817; IMS; 1.
DR PROSITE; PS0173; UMUC; 1.
KW Plasmid; SOS mutagenesis; DNA repair.
FT DOMAIN 2 189 UMUC.
SQ SEQUENCE 424 AA; 47727 MW; FF8C47476CC58A2B CRC64;

Query Match 28.7%; Score 48; DB 1; Length 424;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 10 IRTPDNPANVYAGRGIRP 27
Db 381 INHPGKGIWFAGRGIAP 398

RESULT 12
T2D3 HUMAN
ID T2D3 HUMAN STANDARD; PRT; 1083 AA.
AC Q00268; Q99721; Q9BR40; Q9BX42;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Transcription initiation factor TFIID 135 kDa subunit (TAFII-135)
DE (TAFII135) (TAFII-130) (TAFII130).
GN TAF4 OR TAF4A OR TAF2C1 OR TAF2C OR TAFII135 OR TAFII130.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97336072; PubMed=9192867;
RA Mengus G., May M., Carre L., Chambon P., Davidson I.;
RT "Human TAF(II)135 potentiates transcriptional activation by the AP-2s
RT of the retinoic acid, vitamin D3, and thyroid hormone receptors in
RT mammalian cells.";
RL Genes Dev. 11:1381-1395(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Clee C.M.,
RA Chapman J., Copley V.E., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Copley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., Fraser L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hamblin S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Levaeslaio M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McKay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Senra H.X., Showkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilmings L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN [3]
RP SEQUENCE OF 105-1083 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=97098442; PubMed=8942982;
RA Tanese N., Saluja D., Vassallo M.F., Chen J.-L., Admon A.;
RT "Molecular cloning and analysis of two subunits of the human TFIID
RT complex: hTAFII130 and hTAFII100.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:13611-13616(1996).
RN [4]
RP IDENTIFICATION IN THE TFC-HAT COMPLEX WITH TAF5L; TAF6L; TADA3L;
RX SUPT3H; TAF5; TRAP; GCN5L2 AND TAF10.
RX MEDLINE=99303588; PubMed=10373431;
RA Brand M., Yamamoto K., Staub A., Tora L.;
RT "Identification of TAF4-binding protein-free TAFII-containing complex
RT subunits suggests a role in nucleosome acetylation and signal
RT transduction.";
RL J. Biol. Chem. 274:18285-18289(1999).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 870-918 IN COMPLEX WITH
RX TAF12.
RX MEDLINE=20063193; PubMed=10594036;
RX Gangloff Y.-G., Werten S., Romier C., Carre L., Poch O., Moras D.,
```

Davidson I.
RA "The human TFIIID components TAF(II)135 and TAF(II)20 and the yeast
RT SAGA components ADA1 and TAF(II)68 heterodimerize to form histone-like
RT pairs." ;
RL Mol. Cell. Biol. 20:340-351(2000).
CC -!- FUNCTION: Makes part of TFIIID is a multimeric protein complex that
CC plays a central role in mediating promoter responses to various
CC activators and repressors. Potentiates transcriptional activation
CC by the AP-2s of the retinoic acid, vitamin D3 and thyroid hormone.
CC -!- SUBUNIT: TFIIID is composed of TATA binding protein (TBP) and a
CC number of TBP-associated factors (TAFs). Component of the TFIIC-HAT
CC complex, at least composed of TAF5L, TAF6L, TADA3L, SUPT3H/SPT3,
CC TAF2/TFII150, TAF4/TFII135, TAF5/TFII100, GCN5L2/GCN5, TAF10
CC and TRRAP.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE TAF2C FAMILY.
CC -!- SIMILARITY: Contains 1 TAFH/NHR1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y11354; CAA72189.1; -;
CC DR EMBL: AL137077; CAC36006.1; -;
CC DR EMBL: AL109911; CAC22312.2; -;
CC DR EMBL: U75308; AAC50901.1; -;
CC DR PDB: 1H30; 26-SEP-02.
CC DR TRANSFAC: T02328; -;
CC DR Genew: HGNC:11537; TAF4.
CC DR MIM: 601796; -;
CC DR GO: GO:0005669; C:transcription factor TFIIID complex; TAS.
CC DR GO: GO:0016251; F:general RNA polymerase II transcription fac. .; TAS.
CC DR GO: GO:0005515; F:protein binding; TAS.
CC DR GO: GO:0003713; F:transcription co-activator activity; TAS.
CC DR InterPro: IPR007900; TAF4.
CC DR InterPro: IPR003894; TAF_hom.
CC DR Pfam: PF05236; TAF4; 1.
CC DR SMART: SM00549; TAFH; 1.
CC KW Transcription regulation; Nuclear protein; 3D-structure.
FT DOMAIN 590 682 TAFH/NHR1.
FT DOMAIN 39 42 POLY-HIS.
FT DOMAIN 52 57 POLY-ALA.
FT DOMAIN 98 101 POLY-GLY.
FT DOMAIN 142 148 POLY-ALA.
FT DOMAIN 268 275 POLY-PRO.
FT DOMAIN 331 337 POLY-ALA.
FT DOMAIN 680 683 POLY-PRO.
FT DOMAIN 808 813 POLY-ALA.
FT DOMAIN 828 831 POLY-ASP.
FT CONFLICT 105 117 PGPPSPRPPLVPA -> GRGLLQRRGGRS
(IN REF. 3).
FT CONFLICT 136 136 A -> S (IN REF. 2).
FT CONFLICT 185 185 G -> GPG (IN REF. 2).
FT CONFLICT 233 264 MISSING (IN REF. 3).
FT CONFLICT 293 293 P -> L (IN REF. 3).
SQ SEQUENCE 1083 AA; 109943 MW; A6453827572A0752 CRC64;

Query Match 28.4%; Score 47.5; DB 1; Length 1083;
Best Local Similarity 30.8%; Pred. No. 53;
Matches 12; Conservative 4; Mismatches 8; Indels 15; Gaps 1;

QY 1 SRAHQSMIEIRTPDI-----NPAWVAGRG 24
DB 36 SAAHHHLAPRTPEVRAAAGALGNHWSGSPAGAAG 74

RESULT 13
XERC_VIBVU
ID XERC_VIBVU STANDARD; PRT; 316 AA.

Q7ZA19;
AC 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tyrosine recombinase xerc.
GN XERC OR VV11129.
OS Vibrio vulnificus.
CC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
CC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RT Choy H.E.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Site-specific tyrosine recombinase, which acts by
CC catalyzing the cutting and rejoining of the recombinating DNA
CC molecules. The xerc-xerc complex is essential to convert dimers of
CC the bacterial chromosome into monomers to permit their segregation
CC at cell division. It also contributes to the segregational
CC stability of plasmids (By similarity)
CC -!- SUBUNIT: Forms a cyclic heterotetrameric complex composed of two
CC molecules of xerc and two molecules of xerd (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the "phage" integrase family. Xerc
CC subfamily 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB016800; AAC09604.1; -;
CC DR HAVAP; MF_G1808; -; 1.
CC DR InterPro: IPR004107; Phage integr N.
CC DR InterPro: IPR002104; Phage integrase.
CC DR Pfam: PF02899; Phage integr N; 1.
CC DR Pfam: PF00589; Phage integrase; 1.
CC KW DNA recombination; DNA integration; Cell division;
FT ACT_SITE 157 157 Chromosome partition; DNA-binding; Complete proteome.
FT ACT_SITE 181 181 BY SIMILARITY.
FT ACT_SITE 250 250 BY SIMILARITY.
FT ACT_SITE 253 253 BY SIMILARITY.
FT ACT_SITE 276 276 BY SIMILARITY.
FT ACT_SITE 285 285 TRANSIENT COVALENT LINKAGE TO DNA DURING
STRAND CLEAVAGE AND REJOINING (BY
SIMILARITY).
SQ SEQUENCE 316 AA; 35712 MW; 7728CEB2C968CABC CRC64;

Query Match 28.1%; Score 47; DB 1; Length 316;
Best Local Similarity 37.0%; Pred. No. 18;
Matches 10; Conservative 6; Mismatches 7; Indels 4; Gaps 2;

QY 4 HQHSMEIRT-PDINPAW---YAGSGIR 26
DB 48 HLHSMGLKAWPQVDAGWVRLAGKGM 74

RESULT 14
DH12_RABIT
ID DH12_RABIT STANDARD; PRT; 406 AA.
AC P51976;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Corticosteroid 11-beta-dehydrogenase, isozyme 2 (EC 1.1.1.146) (11-
DH2) (11-beta-hydroxysteroid dehydrogenase type 2) (11-beta-HSD2)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 65.7849 Seconds
(without alignments)
133.146 Million cell updates/sec

Title: US-09-700-643A-1
Perfect score: 167
Sequence: 1 SRASHSMEIRTPDINPAWYAGRIRPVGRX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1980s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	166	99.4	31	2 AAW31371	Aaw31371 Bovine G
2	166	99.4	31	2 AAW95188	Aaw95188 Bovine pi
3	166	99.4	31	2 AAW97613	Aaw97613 Bovine 19
4	166	99.4	31	2 AAW97218	Aaw97218 Bovine pi
5	166	99.4	31	3 AAY49288	Aay49288 19P2 liga
6	166	99.4	31	3 AAY49290	Aay49290 19P2 liga
7	166	99.4	31	3 AAB10347	Aab10347 Bovine ox
8	166	99.4	31	4 AAG62516	Aag62516 Bovine CR
9	166	99.4	31	5 AAE26399	Aae26399 Bovine Pr
10	166	99.4	31	6 ABUS0825	Abus0825 Peptide p
11	166	99.4	31	6 ABUS0831	Abus0831 Peptide p
12	166	99.4	32	2 AAW31372	Aaw31372 Bovine G
13	166	99.4	32	2 AAW95189	Aaw95189 Bovine pi
14	166	99.4	32	3 AAB10348	Aab10348 Bovine ox
15	166	99.4	32	4 AAG62517	Aag62517 Bovine CR
16	166	99.4	32	6 ABUS0832	Abus0832 Peptide p
17	166	99.4	33	2 AAW31373	Aaw31373 Bovine G
18	166	99.4	33	2 AAW95190	Aaw95190 Bovine pi
19	166	99.4	33	3 AAY49297	Aay49297 19P2 liga
20	166	99.4	33	3 AAB10349	Aab10349 Bovine ox
21	166	99.4	33	4 AAG62518	Aag62518 Bovine CR
22	166	99.4	33	6 ABUS0833	Abus0833 Peptide p
23	166	99.4	98	2 AAW31382	Aaw31382 Bovine ge
24	166	99.4	98	2 AAW31368	Aaw31368 Bovine G
25	166	99.4	98	2 AAW95187	Aaw95187 Bovine ge

26	166	99.4	98	2 AAW97217	Aaw97217 Bovine pi
27	166	99.4	98	2 AAW97224	Aaw97224 Bovine ge
28	166	99.4	98	3 AAB10353	Aab10353 Bovine ox
29	166	99.4	98	3 AAB10346	Aab10346 Bovine ox
30	166	99.4	98	4 AAG62522	Aag62522 Bovine CR
31	166	99.4	98	4 AAG62515	Aag62515 Bovine CR
32	166	96.4	29	2 AAW31369	Aaw31369 Bovine G
33	161	96.4	29	2 AAW95184	Aaw95184 Bovine pi
34	161	96.4	29	6 ABUS0829	Abus0829 Peptide p
35	160	95.8	31	4 AAB90995	Aab90995 Prolactin
36	158	94.6	31	4 AAB73370	Aab73370 bPrRp31 p
37	157	94.0	31	2 AAW31384	Aaw31384 Rat type
38	157	94.0	31	2 AAW95174	Aaw95174 Murine pi
39	157	94.0	31	2 AAW95173	Aaw95173 Murine pi
40	157	94.0	31	2 AAW97614	Aaw97614 Rat 19P2
41	157	94.0	31	2 AAW97233	Aaw97233 Rat type
42	157	94.0	31	3 AAY49292	Aay49292 19P2 liga
43	157	94.0	31	3 AAY87504	Aay87504 Rat prola
44	157	94.0	31	3 AAB10355	Aab10355 Rat oxyto
45	157	94.0	31	4 AAB90993	Aab90993 Prolactin

ALIGNMENTS

RESULT 1
AAW31371
ID AAW31371 standard; peptide; 31 AA.

XX AAW31371;

XX 06-APR-1998 (first entry)

XX Bovine G protein-coupled receptor ligand peptide fragment 1.

XX G protein-coupled receptor; ligand binding; pharmaceutical; modulator;
XX pituitary; central nervous system; pancreas; prophylactic;
XX therapeutic agent.

XX Bos taurus

XX WO9724436-A2.

XX 10-JUL-1997.

XX 26-DEC-1996; 96WO-JP003821.

XX 28-DEC-1995; 95JP-00343371.

XX 15-MAR-1996; 96JP-00059419.

XX 12-AUG-1996; 96JP-00211805.

XX 18-SEP-1996; 96JP-00246573.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;

XX Kitada C;

XX WPI; 1997-363672/33.

XX N-PSDB; AAV02394.

XX Ligand peptide for G protein-coupled receptor - acts by modulating

XX function in the central nervous system, pancreas and pituitary gland.

XX Claim 2; Page 160; 258pp; English.

XX This sequence represents a peptide fragment of a novel bovine pituitary
XX derived ligand corresponding to amino acid residues 23 to 53 of the
XX sequence in AAW31368 and is used in an assay to monitor ligand binding to
XX the G protein-coupled receptor protein. Pharmaceutical compositions
XX containing this ligand may be used as a pituitary function modulator, a
XX central nervous system modulator or a pancreatic function modulator. This
XX ligand could have specific applications as a prophylactic or therapeutic
XX agent for dementia, depression, hyperkinetic syndrome, disturbance of

CC consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone
 CC secretory disease, hyper- and polyphagia, hypercholesterolaemia,
 CC hyperglycaemia, hyperlipidaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia,
 CC amyotrophic lateral sclerosis, acute myocardial infarction,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis, asthma, epilepsy, infertility and/or oligogalactia. Assays
 CC can also be developed to screen compounds which are capable of altering
 CC the binding activity of the ligand thus affecting activation of the G
 CC protein-coupled receptor protein
 XX
 XX Sequence 31 AA;

Query Match 99.4%; Score 166; DB 2; Length 31;

Best Local Similarity 100.0%; Pred. NO. 9.2e-18;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAHQHSEIRTPDINPAWAGRGIRPVGR 30

DB 1 SRAHQHSEIRTPDINPAWAGRGIRPVGR 30

RESULT 2

AAW95188
 ID AAW95188 standard; peptide; 31 AA.

AC AAW95188;

DT 10-MAR-1999 (first entry)

XX Bovine pituitary-derived ligand polypeptide fragment.

XX Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
 KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
 KW Creutzfeldt-Jakob disease; poisoning; schizophrenia; growth hormone;
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
 KW gene therapy; transgenic animal; bovine.

XX BOS SP.

XX WO9849295-A1.

XX 05-NOV-1998.

XX 27-APR-1998; 98WO-JF001923.

XX 28-APR-1997; 97JP-00109974.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Hinuma S, Fukusumi S;

XX WPI; 1999-009423/01.

XX New polypeptide ligand for orphan G protein coupled receptors - used for
 PT treating disorders of central nervous system, pituitary and pancreas, and
 PT for drug screening.

XX Example 19; Page 150; 206pp; English.

XX The invention relates to a murine pituitary-derived ligand polypeptide
 CC which is a ligand for the G-protein coupled orphan receptor designated
 CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing
 CC the ligand polypeptide encoding DNA are used to produce a recombinant
 CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate
 CC function of the pituitary, central nervous system, pancreas and other
 CC tissues and can be used to screen for agents that modulate binding of the
 CC polypeptide to the receptor; to quantify the amount of receptor in a
 CC sample and to raise antibodies. They may also be used therapeutically,
 CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's

CC diseases; Creutzfeldt-Jakob disease; poisoning by heavy metals or drugs;
 CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;
 CC rheumatoid arthritis, epilepsy and many others, also to improve post-
 CC operative nutritional status and as vasopressor. Transgenic animals
 CC carrying the ligand polypeptide encoding DNA or its mutain are used to
 CC study the function of the polypeptide-expressing genes, as models of
 CC disease, for drug screening and as source of cell lines. The ligand
 CC polypeptide DNA is used as a source of probes and primers; to identify
 CC related sequences; in receptor-binding assays; for production of Ab and
 CC antisera; in drug development; for gene therapy and to develop transgenic
 CC animals. The present sequence represents a bovine genome-derived ligand
 CC polypeptide fragment which is similar to the murine ligand-polypeptide
 XX
 XX Sequence 31 AA;

Query Match 99.4%; Score 166; DB 2; Length 31;

Best Local Similarity 100.0%; Pred. NO. 9.2e-18;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAHQHSEIRTPDINPAWAGRGIRPVGR 30

DB 1 SRAHQHSEIRTPDINPAWAGRGIRPVGR 30

RESULT 3

AAW87613
 ID AAW87613 standard; peptide; 31 AA.

AC AAW87613;

DT 29-MAR-1999 (first entry)

XX Bovine 19P2 ligand.

XX 19P2 ligand; G protein coupled receptor; pituitary;
 KW prolactin releasing peptide; cattle; dementia; breast cancer; therapy.

XX Bos taurus.

XX EP887417-A2.

XX 30-DEC-1998.

XX 25-JUN-1998; 98EP-00111725.

XX 27-JUN-1997; 97JP-00172118.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Suenaga M, Moriya T, Tanaka Y, Nishimura O;

XX WPI; 1999-047884/05.

XX Producing a 19P2 pituitary G protein receptor ligand - by cleavage of a
 PT fusion protein, useful for preventing and treating dementia, breast
 PT cancer, renal failure and autoimmune disease.

XX Claim 5; Page 34; 56pp; English.

XX This is the amino acid sequence of the bovine pituitary G protein-coupled
 CC receptor ligand 19P2L. A method suitable for commercial high-level
 CC production of 19P2L comprises expressing the ligand in host cells as a
 CC recombinant fusion protein e.g. with human basic fibroblast growth factor
 CC (see AAV83792-93) that has been modified to include an N-terminal
 CC cysteine residue. The ligand is released from the fusion by cyanylation
 CC followed by ammonolysis. 19P2L has prolactin secretion-stimulating and
 CC (at high doses) prolactin secretion-inhibiting properties. It can be used
 CC in the treatment and prevention of various diseases including: senile
 CC dementia, cerebrovascular dementia, and dementia associated with:
 CC Genealogical disorders (e.g. Alzheimer's disease, Parkinson's disease,
 CC Pick's disease, Huntington's disease), infectious diseases (e.g.
 CC Creutzfeldt-Jakob's), endocrine or metabolic disease or toxicosis (e.g.
 CC hypothyroidism, vitamin B12 deficiency, alcoholism, intoxication by

CC drugs, metal and organic compounds), tumorigenic diseases (e.g. brain
CC tumour), traumatic diseases (e.g. chronic subarachnoidal hemorrhage, and
CC other types of dementia, depression, hyperactive child syndrome
CC (microcephalopathy) and disturbance of consciousness. It is also useful
CC for prevention and treatment of diseases associated with prolactin hypo
CC and hypersecretion respectively, including: hyperprolactinaemia,
CC pituitary adenoma, breast cancer, infertility, impotence and autoimmune
CC disease (hypersecretion disorders), and seminal vesicle hypoplasia,
CC osteoporosis, menopausal syndrome and renal failure (hyposecretion
CC disorders). The 19p2 polypeptide/amide is also useful as a test reagent
CC for study of the prolactin secretory function or as a lactogogue in
CC mammalian farm animals
XX
SQ Sequence 31 AA;
Query Match 99.4%; Score 166; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 9.2e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQSMSEIRTPDINPAWYAGRGIRPVGR 30
DB 1 SRAHQSMSEIRTPDINPAWYAGRGIRPVGR 30
RESULT 4
AAW97218
ID AAW97218 standard; peptide; 31 AA.
AC
XX
XX AAW97218;
DT 06-MAY-1999 (first entry)
DE Bovine pituitary-derived ligand polypeptide fragment.
XX
XX Bovine pituitary-derived ligand; modulation; prolactin secretion;
KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;
KW menopausal syndrome; euthyroid; hypometabolism; lactation;
KW pituitary adenomatosis; brain tumour; emmeniopathy; autoimmune disease;
KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;
KW acromegaly; Chiari-Frömmel syndrome; Argonz-del Castillo syndrome;
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;
KW contraceptive; placental function; choriocarcinoma; hydatid mole;
KW interruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;
KW abnormal lipidmetabolism; oxytocia.
XX
XX Bos SP.
XX WO9858962-A1.
XX
XX 30-DEC-1998.
XX
XX 22-JUN-1998; 98WO-JP002765.
XX
XX 23-JUN-1997; 97JP-00165437.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Hinuma S, Kawamata Y, Fujii R, Matsumoto H;
XX
XX WPI; 1999-105614/09.
XX
XX Use of G protein-coupled receptor ligands - for modulating prolactin
PT secretion or placental function, e.g. for treating menopausal syndrome,
PT tumours, autoimmune disease or abnormal pregnancy.
XX
XX Claim 3; Page 135; 241pp; English.
XX
XX The present sequence represents a bovine pituitary-derived ligand
CC fragment. It is used in the course of the invention. The specification
CC describes an agent for modulating prolactin secretion which comprises a
CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)
CC protein. The agents for promoting prolactin secretion can be used for
CC treating or preventing hypovarianism, gonocyst cacogenesis, menopausal

CC syndrome, euthyroid or hypometabolism. They can be used for promoting
CC lactation in a domestic mammal and as an aphrodisiac. The agents for
CC inhibiting prolactin secretion can be used for treating or preventing
CC pituitary adenomatosis, brain tumour, emmeniopathy, autoimmune disease,
CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,
CC acromegaly, Chiari-Frömmel syndrome, Argonz-del Castillo syndrome, Forbes-
CC Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia. The
CC inhibitory agents can also be used as contraceptives. The agents for
CC modulating placental function can be used for treating or preventing
CC choriocarcinoma, hydatid mole, interruption mole, abortion, unthrifty fetus,
CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia
XX
SQ Sequence 31 AA;
Query Match 99.4%; Score 166; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 9.2e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQSMSEIRTPDINPAWYAGRGIRPVGR 30
DB 1 SRAHQSMSEIRTPDINPAWYAGRGIRPVGR 30
RESULT 5
AAW49298
ID AAW49298 standard; peptide; 31 AA.
XX
XX AAW49298;
XX
XX 22-FEB-2000 (first entry)
XX
XX 19p2 ligand peptide fragment.
XX Monoclonal antibody; 19p2 ligand; diagnosis; prolactin secretion;
KW pituitary; regulatory mechanism; central nervous system; pancreatic.
XX Homo sapiens.
XX Key Location/Qualifiers
FH Modified-site 31
FT /note= "C-terminal amide"
XX
XX WO9960112-A1.
XX
XX 25-NOV-1999.
XX
XX 20-MAY-1999; 99WO-JP002650.
XX
XX 21-MAY-1998; 98JP-00140293.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Matsumoto H, Kitada C, Hinuma S;
XX
XX WPI; 2000-039381/03.
XX
XX New monoclonal antibodies, useful in diagnosis, as drugs and in studying
PT diseases related to ligand abnormality.
XX
XX Disclosure; Page 27; 73pp; Japanese.
XX
XX The invention provides a monoclonal antibody which has a specific
CC reaction with the part peptide of the C-terminal of 19p2 ligand or its
CC derivative. The antibodies can be used in diagnosis or to treat or
CC prevent diseases associated with abnormality in the pituitary function
CC regulatory mechanism (e.g. promotion of prolactin secretion), central
CC nervous regulatory mechanism, and pancreatic function regulatory
CC mechanism. The antibody-based immunoassay can also be applied in
CC clarifying the physiological functions of the ligand and its derivative.
CC Sequences AAW49298-302 represent peptide fragments of the 19p2 ligand
XX
SQ Sequence 31 AA;

Query Match 99.4%; Score 166; DB 3; Length 31;
 Best Local Similarity 100.0%; Pred. No. 9.2e-18;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAHQHSEIRTPDINPAWYAGRGIRPVGR 30
 DB 1 SRAHQHSEIRTPDINPAWYAGRGIRPVGR 30

RESULT 6
 AAY49290
 ID AAY49290 standard; peptide; 31 AA.
 AC AAY49290;
 XX
 DT 22-FEB-2000 (first entry)
 XX
 DE 19P2 ligand peptide fragment.
 XX
 KW Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;
 KW pituitary; regulatory mechanism; central nervous system; pancreatic.
 XX
 OS Bos sp.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 31 /note= "C-terminal amide"
 FT WO9960112-A1.
 XX
 XX 25-NOV-1999.
 XX
 XX 20-MAY-1999; 99WO-JP002650.
 XX
 XX 21-MAY-1998; 98JP-00140293.
 XX
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX
 XX Matsumoto H, Kitada C, Hinuma S;
 XX WPI; 2000-039381/03.
 XX
 XX New monoclonal antibodies, useful in diagnosis, as drugs and in studying
 PT diseases related to ligand abnormality.
 XX
 XX Disclosure; Page 26; 73pp; Japanese.
 PS
 CC The invention provides a monoclonal antibody which has a specific
 CC reaction with the part peptide of the C-terminal of 19P2 ligand or its
 CC derivative. The antibodies can be used in diagnosis or to treat or
 CC prevent diseases associated with abnormality in the pituitary function
 CC regulatory mechanism (e.g. promotion of prolactin secretion), central
 CC nervous regulatory mechanism, and pancreatic function regulatory
 CC mechanism. The antibody-based immunoassay can also be applied in
 CC clarifying the physiological functions of the ligand and its derivative.
 CC Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand
 XX
 XX Sequence 31 AA;

Query Match 99.4%; Score 166; DB 3; Length 31;
 Best Local Similarity 100.0%; Pred. No. 9.2e-18;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAHQHSEIRTPDINPAWYAGRGIRPVGR 30
 DB 1 SRAHQHSEIRTPDINPAWYAGRGIRPVGR 30

RESULT 7
 AAB10347
 ID AAB10347 standard; peptide; 31 AA.
 XX
 AC AAB10347;

XX
 DT 24-NOV-2000 (first entry)
 XX
 DE Bovine oxytocin secretion promoting peptide SEQ ID NO: 3.
 XX
 KW Bovine; oxytocin secretion promoter; G protein-coupled receptor protein;
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;
 KW veterinary medicine; milk production.
 XX
 OS Bos taurus.
 XX
 PN WO200038704-A1.
 XX
 PD 06-JUL-2000.
 XX
 PF 22-DEC-1999; 99WO-JP007199.
 XX
 PR 25-DEC-1998; 98JP-00369585.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Matsumoto H, Kitada C, Hinuma S;
 XX WPI; 2000-452298/39.
 XX
 PT Physiologically-active polypeptide recognized as ligand by G protein-
 PT coupled receptor protein, for promoting secretion of oxytocin, as drugs
 PT for diseases relating to oxytocin secretion and in veterinary medicine.
 XX
 PS Claim 3; Page 50; 72pp; Japanese.
 XX
 CC This invention describes a novel oxytocin secretion-regulating agent
 CC which contains a ligand peptide or its salt for the G protein-coupled
 CC receptor protein. It is useful in the form of drugs for ameliorating,
 CC preventing and treating diseases relating to oxytocin secretion e.g. weak
 CC pains and atonic bleeding, before and after expulsion of placenta,
 CC uterine recovery failure, caesarean section, stoppage of artificial
 CC fertilization or galactostasis and is also applicable in veterinary
 CC medicine for promoting milk production in cow, goat and pig. This
 CC sequence represents a bovine peptide which acts as an oxytocin secretion
 CC promoter
 XX
 XX Sequence 31 AA;

Query Match 99.4%; Score 166; DB 3; Length 31;
 Best Local Similarity 100.0%; Pred. No. 9.2e-18;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAHQHSEIRTPDINPAWYAGRGIRPVGR 30
 DB 1 SRAHQHSEIRTPDINPAWYAGRGIRPVGR 30

RESULT 8
 AAG62516
 ID AAG62516 standard; peptide; 31 AA.
 XX
 AC AAG62516;
 XX
 DT 24-AUG-2001 (first entry)
 XX
 DE Bovine CRH releasing protein related peptide SEQ ID NO: 3.
 XX
 KW Cow; corticotrophin releasing hormone; CRH; G protein receptor ligand;
 KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;
 KW Addison's disease; adrenal gland hyperfunction; obesity.
 XX
 OS Bos taurus.
 XX
 PN WO200135984-A1.
 XX
 XX 25-MAY-2001.

XX 17-NOV-2000; 2000WO-JP008119.
 XX 18-NOV-1999; 99JP-00327900.
 PR 26-SEP-2000; 2000JP-00297073.
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX Kitada C, Matsumoto H, Hinuma S;
 PI WPI; 2001-355552/37.
 DR
 XX Use of G protein receptor ligand or peptide for controlling corticotropin
 PT releasing hormone secretion.
 XX
 PS Claim 3; Page 63; 90pp; Japanese.
 XX The present sequence describes a method of controlling the secretion of
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein
 CC receptor ligand. This can be used to control the secretion of CRH and is
 CC useful as an analgesic or for treating, preventing or ameliorating
 CC diseases associated with CRH secretion such as hyperaldosteronism,
 CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The
 CC present sequence is a peptide used in the exemplification of the
 CC invention
 XX
 SQ Sequence 31 AA;
 Query Match 99.4%; Score 166; DB 4; Length 31;
 Best Local Similarity 100.0%; Pred. No. 9.2e-18;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGR 30
 DB 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGR 30
 RESULT 9
 AAE26399
 ID AAE26399 standard; peptide; 31 AA.
 XX
 AC AAE26399;
 XX
 DT 13-DEC-2002 (first entry)
 XX
 DE Bovine PrRP-31 peptide.
 XX
 KW Wakefulness; sleep disorder; prolactin releasing peptide receptor; PrRP;
 KW GPR10; therapy; epilepsy; narcolepsy; sleepiness; sleep apnoea; insomnia;
 KW idiopathic hypersomnia; psychogenic hypersomnia; seizure; anticonvulsant;
 KW bovine.
 XX
 CS Bos taurus.
 XX
 XX US2002037533-A1.
 XX
 XX 28-MAR-2002.
 XX
 XX 17-AUG-2001; 2001US-00932161.
 XX
 XX 28-APR-2000; 2000US-00560915.
 XX
 XX (CIVE/) CIVELLI O.
 XX (LINS/) LIN S.
 PA
 PA Civelli O, Lin S;
 PI
 XX
 XX WPI; 2002-403991/43.
 DR
 XX Screening for compounds useful for promoting wakefulness or sleep, and
 PT for treating sleeping disorders, e.g. insomnia, hypersomnia or sleep

PT apnea, comprises administering a prolactin releasing peptide agonist or
 PT antagonist.
 XX
 PS Disclosure; Page 24; 35pp; English.
 XX
 CC The present invention relates to a method of screening for compounds for
 CC promoting wakefulness or sleep in a mammal. The method involves
 CC administering a prolactin releasing peptide (PrRP) receptor (GPR10)
 CC agonist or antagonist respectively and determining the ability of the
 CC compound to promote wakefulness or sleep. The compounds identified from
 CC the method are used in the therapy of epilepsy and other diseases
 CC associated with absence seizures and in promoting wakefulness and sleep
 CC in individuals having sleep disorders such as insomnia and narcolepsy.
 CC PrRP receptor agonists may be used to treat common disorders which lead
 CC to sleepiness, e.g. sleep apnoea, narcolepsy, idiopathic hypersomnia and
 CC psychogenic hypersomnia. PrRP receptor antagonists are useful for
 CC promoting sleep and for treating insomnia such as adjustment sleep
 CC disorder and psychophysiological insomnia. The present sequence is bovine
 CC PrRP-31 peptide
 XX
 SQ Sequence 31 AA;
 Query Match 99.4%; Score 166; DB 5; Length 31;
 Best Local Similarity 100.0%; Pred. No. 9.2e-18;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGR 30
 DB 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGR 30
 RESULT 10
 ABU60825
 ID ABU60825 standard; peptide; 31 AA.
 XX
 AC ABU60825;
 XX
 DT 06-MAY-2003 (first entry)
 XX
 DE Peptide production by gene recombination associated peptide #9.
 XX
 KW Peptide production; low-molecular peptide; KiSS-1; GPR8 ligand;
 KW gene recombination.
 XX
 OS Bos sp.
 XX
 XX WO200292829-A1.
 XX
 XX 21-NOV-2002.
 XX
 XX 16-MAY-2002; 2002WO-JP004735.
 XX
 XX 17-MAY-2001; 2001JP-00147341.
 XX
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX
 XX Nishimura O, Suenaga M, Ito T, Kitada C;
 PI
 XX WPI; 2003-129302/12.
 XX
 XX Process for producing peptides e.g. KiSS-1 peptide and GPR8 ligand for
 XX subsequent applications by gene recombination technique through tandem
 XX repeats to provide precursor protein with specific cleavage sites.
 PT
 XX Disclosure; Page 58; 87pp; Japanese.
 PS
 XX The invention describes a method of producing a peptide comprising the
 XX excision of the N and C-terminals of a target peptide with enzymes or
 XX chemically through the attached cleavage sites repeated by ligation in a
 XX precursor protein. The method is for producing (low-molecular) peptides
 XX e.g. KiSS-1 peptide and GPR8 ligand for subsequent applications by the
 XX gene recombination technique through tandem repeats to provide a
 XX precursor protein with specific cleavage sites. With this method, peptide

CC production can be carried out easily to provide large quantities of the
 CC required peptides. This is the amino acid sequence of a peptide
 CC associated with the peptide production method of the invention

XX SQ Sequence 31 AA;

Query Match 99.4%; Score 166; DB 6; Length 31;
 Best Local Similarity 100.0%; Pred. No. 9.2e-18;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETPDINPAWYAGRGIRPVGR 30
 |||||
 Db 1 SRAHQHSMETPDINPAWYAGRGIRPVGR 30

RESULT 11

ABU60831
 ID ABU60831 standard; peptide; 31 AA.

XX AC ABU60831;

XX DT 06-MAY-2003 (first entry)

XX DE Peptide production by gene recombination associated peptide #15.

XX KW Peptide production; low-molecular peptide; KiSS-1; GPR8 ligand;

XX KW gene recombination.

XX OS Bos sp.

XX PN WO200292829-A1.

XX PD 21-NOV-2002.

XX PF 16-MAY-2002; 2002WO-JP004735.

XX PR 17-MAY-2001; 2001JP-00147341.

XX PA (TAKE) TAKEDA CHEM IND LTD.

XX PI Nishimura O, Suenaga M, Ito T, Kitada C;

XX DR WPI; 2003-129302/12.

XX PT Process for producing peptides e.g. KiSS-1 peptide and GPR8 ligand for
 PT subsequent applications by gene recombination technique through tandem
 PT repeats to provide precursor protein with specific cleavage sites.

XX PS Disclosure; Page 61; 87pp; Japanese.

XX CC The invention describes a method of producing a peptide comprising the
 CC excision of the N and C-terminals of a target peptide with enzymes or
 CC chemically through the attached cleavage sites repeated by ligation in a
 CC precursor protein. The method is for producing (low-molecular) peptides
 CC e.g. KiSS-1 peptide and GPR8 ligand for subsequent applications by the
 CC gene recombination technique through tandem repeats to provide a
 CC precursor protein with specific cleavage sites. With this method, peptide
 CC production can be carried out easily to provide large quantities of the
 CC required peptides. This is the amino acid sequence of a peptide
 CC associated with the peptide production method of the invention

XX SQ Sequence 31 AA;

Query Match 99.4%; Score 166; DB 6; Length 31;
 Best Local Similarity 100.0%; Pred. No. 9.2e-18;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETPDINPAWYAGRGIRPVGR 30
 |||||

Db 1 SRAHQHSMETPDINPAWYAGRGIRPVGR 30
 |||||

RESULT 12

AAW31372

ID AAW31372 standard; peptide; 32 AA.

XX AC AAW31372;

XX DT 06-APR-1998 (first entry)

XX DE Bovine G protein-coupled receptor ligand peptide fragment 2.

XX KW G protein-coupled receptor; ligand binding; pharmaceutical; modulator;
 KW pituitary; central nervous system; pancreas; prophylactic;
 KW therapeutic agent.

XX OS Bos taurus.

XX PN WO9724436-A2.

XX PD 10-JUL-1997.

XX PF 26-DEC-1996; 96WO-JP003821.

XX PR 28-DEC-1995; 95JP-00343371.

XX PR 15-MAR-1996; 96JP-00059419.

XX PR 12-AUG-1996; 96JP-00211805.

XX PR 18-SEP-1996; 96JP-00246573.

XX PA (TAKE) TAKEDA CHEM IND LTD.

XX PI Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;

XX PI Kitada C;

XX DR WPI; 1997-363672/33.

XX DR N-PSDB; AAV02395.

XX PT Ligand peptide for G protein-coupled receptor - acts by modulating
 PT function in the central nervous system, pancreas and pituitary gland.

XX PS Claim 2; Page 160-161; 258pp; English.

XX CC This sequence represents a peptide fragment of a novel bovine pituitary
 CC derived ligand corresponding to amino acid residues 23 to 54 of the
 CC sequence in AAW31368 and is used in an assay to monitor ligand binding to
 CC the G protein-coupled receptor protein. Pharmaceutical compositions
 CC containing this ligand may be used as a pituitary function modulator, a
 CC central nervous system modulator or a pancreatic function modulator. This
 CC ligand could have specific applications as a prophylactic or therapeutic
 CC agent for dementia, depression, hyperkinetic syndrome, disturbance of
 CC consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone
 CC secretory disease, hyper- and polyphagia, hypercholesterolaemia,
 CC hyperglycaemia, hyperlipidaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia,
 CC amyotrophic lateral sclerosis, acute myocardial infarction,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis, asthma, epilepsy, infertility and/or oligogalactia. Assays
 CC can also be developed to screen compounds which are capable of altering
 CC the binding activity of the ligand thus affecting activation of the G
 CC protein-coupled receptor protein

XX SQ Sequence 32 AA;

Query Match 99.4%; Score 166; DB 2; Length 32;
 Best Local Similarity 100.0%; Pred. No. 9.6e-18;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETPDINPAWYAGRGIRPVGR 30
 |||||

Db 1 SRAHQHSMETPDINPAWYAGRGIRPVGR 30
 |||||

RESULT 13

AAW95189

ID AAW95189 standard; peptide; 32 AA.

XX AAW95189;
 AC 10-MAR-1999 (first entry)
 DT
 XX
 DE Bovine pituitary-derived ligand polypeptide fragment.
 XX
 KW Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
 KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
 KW Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
 KW gene therapy; transgenic animal; Bovine.
 XX
 OS Bos sp.
 XX
 XX WO9849295-A1.
 PN
 PD 05-NOV-1998.
 XX
 XX 27-APR-1998; 98WO-JP001923.
 PF
 XX 28-APR-1997; 97JP-00109974.
 PR
 XX (TAKE) TAKEDA CHEM IND LTD.
 PA
 XX Hinuma S, Fukusumi S;
 PI
 XX WPI; 1999-009423/01.
 DR
 XX
 XX New polypeptide ligand for orphan G protein coupled receptors - used for
 PT treating disorders of central nervous system, pituitary and pancreas, and
 PT for drug screening.
 PT
 XX
 PS Example 19; Page 150; 205pp; English.
 CC
 CC The invention relates to a murine pituitary-derived ligand polypeptide
 CC which is a ligand for the G-protein coupled orphan receptor designated
 CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing
 CC the ligand polypeptide encoding DNA are used to produce a recombinant
 CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate
 CC function of the pituitary, central nervous system, pancreas and other
 CC tissues and can be used to screen for agents that modulate binding of the
 CC polypeptide to the receptor; to quantify the amount of receptor in a
 CC sample and to raise antibodies. They may also be used therapeutically,
 CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's
 CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;
 CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;
 CC rheumatoid arthritis, epilepsy and many others, also to improve post-
 CC operative nutritional status and as vasopressor. Transgenic animals
 CC carrying the ligand polypeptide encoding DNA or its mutin are used to
 CC study the function of the polypeptide-expressing genes, as models of
 CC disease, for drug screening and as source of cell lines. The ligand
 CC polypeptide DNA is used as a source of probes and primers; to identify
 CC related sequences; in receptor-binding assays; for production of Ab and
 CC antisera; in drug development; for gene therapy and to develop transgenic
 CC animals. The present sequence represents a bovine genome-derived ligand
 CC polypeptide fragment which is similar to the murine ligand-polypeptide
 CC
 XX Sequence 32 AA;
 SQ
 Query Match 99.4%; Score 166; DB 2; Length 32;
 Best Local Similarity 100.0%; Pred. No. 9.6e-18;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGR 30
 DB 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGR 30
 RESULT 14
 AAB10348

ID AAB10348 standard; peptide; 32 AA.
 XX
 AC AAB10348;
 XX
 DT 24-NOV-2000 (first entry)
 XX
 DE Bovine oxytocin secretion promoting peptide SEQ ID NO: 4.
 XX
 KW Bovine; oxytocin secretion promoter; G protein-coupled receptor protein;
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;
 KW veterinary medicine; milk production.
 KW
 XX Bos taurus.
 OS
 XX WO200038704-A1.
 PN
 PD 06-JUL-2000.
 XX
 XX 22-DEC-1999; 99WO-JP007199.
 PF
 XX 25-DEC-1998; 98JP-00369585.
 PR
 XX (TAKE) TAKEDA CHEM IND LTD.
 PA
 XX Matsumoto H, Kitada C, Hinuma S;
 PI
 XX WPI; 2000-452298/39.
 DR
 XX Physiologically-active polypeptide recognized as ligand by G protein-
 PT coupled receptor protein, for promoting secretion of oxytocin, as drugs
 PT for diseases relating to oxytocin secretion and in veterinary medicine.
 XX
 PS Disclosure; Page 51; 72pp; Japanese.
 XX
 CC This invention describes a novel oxytocin secretion-regulating agent
 CC which contains a ligand peptide or its salt for the G protein-coupled
 CC receptor protein. It is useful in the form of drugs for ameliorating,
 CC preventing and treating diseases relating to oxytocin secretion e.g. weak
 CC pains and atonic bleeding, before and after expulsion of placenta,
 CC uterine recovery failure, caesarean section, stoppage of artificial
 CC fertilization or galactostasis and is also applicable in veterinary
 CC medicine for promoting milk production in cow, goat and pig. This
 CC sequence represents a bovine peptide which acts as an oxytocin secretion
 CC promoter.
 XX
 SQ Sequence 32 AA;
 Query Match 99.4%; Score 166; DB 3; Length 32;
 Best Local Similarity 100.0%; Pred. No. 9.6e-18;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGR 30
 DB 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGR 30
 RESULT 15
 AAG62517
 ID AAG62517 standard; peptide; 32 AA.
 XX
 AC AAG62517;
 XX
 DT 24-AUG-2001 (first entry)
 XX
 DE Bovine CRH releasing protein related peptide SEQ ID NO: 4.
 XX
 KW Cow; corticotrophin releasing hormone; CRH; G protein receptor ligand;
 KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;
 KW Addison's disease; adrenal gland hyperfunction; obesity.
 KW
 XX Bos taurus.
 OS
 XX

PN WO200135984-A1.
 XX
 PD 25-MAY-2001.
 XX
 PF 17-NOV-2000; 2000WO-JP008119.
 XX
 PR 18-NOV-1999; 99JP-00327900.
 PR 26-SEP-2000; 2000JP-00297073.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 XX
 PI Kitada C, Matsumoto H, Hinuma S;
 XX
 DR WPI; 2001-355552/37.
 XX
 PT Use of G protein receptor ligand or peptide for controlling corticotropin
 PT releasing hormone secretion.
 XX
 PS Disclosure; Page 64; 90pp; Japanese.
 XX
 CC The present sequence describes a method of controlling the secretion of
 CC corticotropin releasing hormone (CRH), involving the use of a G protein
 CC receptor ligand. This can be used to control the secretion of CRH and is
 CC useful as an analgesic or for treating, preventing or ameliorating
 CC diseases associated with CRH secretion such as hyperaldosteronism,
 CC hypercortisolism, secondary or chronic hypoadrenocorticism, Addison's
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The
 CC present sequence is a peptide used in the exemplification of the
 CC invention
 XX
 SQ Sequence 32 AA;

Query Match 99.4%; Score 166; DB 4; Length 32;
 Best Local Similarity 100.0%; Pred. No. 9.6e-18;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRAHQHSMETPTDINPAWYAGRGIRPVGR 30
 |||||
 Db 1 SRAHQHSMETPTDINPAWYAGRGIRPVGR 30

Search completed: August 12, 2004, 14:43:53
 Job time : 65.9099 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:49:10 ; Search time 54.7907 Seconds
(without alignments)
177.617 Million cell updates/sec

Title: US-09-700-643A-1
Perfect score: 167
Sequence: 1 SRAHQSMSEIRTPDINPAWYAGRGIRPVGRX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1292805 segs, 313927144 residues
Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/2/pubpaa/RCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*
- 19: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	166	99.4	31	9	US-09-932-161-13
2	166	99.4	31	13	US-10-044-592-39
3	166	99.4	31	14	US-10-096-777-13
4	166	99.4	32	13	US-10-044-592-40
5	166	99.4	33	13	US-10-044-592-41
6	166	99.4	98	13	US-10-044-592-28
7	166	99.4	98	13	US-10-044-592-38
8	166	99.4	98	13	US-10-044-592-82
9	166	99.4	98	13	US-10-044-592-84
10	166	99.4	98	13	US-10-044-592-86
11	166	99.4	98	13	US-10-044-592-88
12	161	96.4	29	13	US-10-044-592-26
13	157	94.0	31	9	US-09-932-161-14
14	157	94.0	31	13	US-10-044-592-4
15	157	94.0	31	13	US-10-044-592-5

16	157	94.0	31	14	US-10-096-777-14	Sequence 14, Appl
17	157	94.0	70	13	US-10-044-592-90	Sequence 90, Appl
18	157	94.0	82	13	US-10-044-592-1	Sequence 1, Appl
19	157	94.0	86	13	US-10-044-592-96	Sequence 96, Appl
20	157	94.0	91	13	US-10-044-592-94	Sequence 94, Appl
21	152	91.0	31	9	US-09-932-161-15	Sequence 15, Appl
22	152	91.0	31	14	US-10-096-777-15	Sequence 15, Appl
23	152	91.0	87	13	US-10-044-592-92	Sequence 92, Appl
24	142	85.0	25	13	US-10-044-592-78	Sequence 78, Appl
25	109	65.3	19	13	US-10-044-592-27	Sequence 27, Appl
26	109	65.3	20	9	US-09-932-161-16	Sequence 16, Appl
27	109	65.3	20	14	US-10-044-592-42	Sequence 42, Appl
28	109	65.3	20	13	US-10-096-777-16	Sequence 16, Appl
29	109	65.3	21	13	US-10-044-592-43	Sequence 43, Appl
30	109	65.3	22	13	US-10-044-592-44	Sequence 44, Appl
31	105	62.9	20	9	US-09-932-161-17	Sequence 17, Appl
32	105	62.9	20	13	US-10-044-592-6	Sequence 6, Appl
33	105	62.9	20	14	US-10-096-777-17	Sequence 17, Appl
34	103	61.7	20	9	US-09-932-161-18	Sequence 18, Appl
35	103	61.7	20	14	US-10-096-777-18	Sequence 18, Appl
36	96	57.5	40	13	US-10-044-592-80	Sequence 80, Appl
37	61	36.5	428	9	US-09-820-155-2	Sequence 2, Appl
38	57	34.1	9	13	US-10-044-592-8	Sequence 8, Appl
39	57	34.1	465	14	US-10-301-822-197	Sequence 197, App
40	56	33.5	428	9	US-09-820-155-4	Sequence 4, Appl
41	55	32.9	133	12	US-10-424-599-153474	Sequence 153474, A
42	55	32.9	1245	15	US-10-369-493-20447	Sequence 20447, A
43	54	32.3	105	12	US-10-424-599-160071	Sequence 160071, A
44	54	32.3	209	13	US-10-108-915-30	Sequence 30, Appl
45	52	31.1	183	12	US-10-424-599-268092	Sequence 268092, A

ALIGNMENTS

RESULT 1
US-09-932-161-13
; Sequence 13, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Civeili, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; TITLE OF INVENTION: Promoting Wakefulness and Sleep
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932.161
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-932-161-13

Query Match 99.4%; Score 166; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. No. 8.9e-17;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQSMSEIRTPDINPAWYAGRGIRPVGR 30
DB 1 SRAHQSMSEIRTPDINPAWYAGRGIRPVGR 30

RESULT 2
US-10-044-592-39
; Sequence 39, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji


```

; LENGTH: 98
; TYPE: PRT
; ORGANISM: Murine
US-10-044-592-28

Query Match          99.4%; Score 166; DB 13; Length 98;
Best Local Similarity 100.0%; Pred. No. 3e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETPTDINPAWYAGRGIRPVGR 30
    |||||
Db 23 SRAHQHSMETPTDINPAWYAGRGIRPVGR 52

RESULT 7
US-10-044-592-38
; Sequence 38, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 38
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-38

Query Match          99.4%; Score 166; DB 13; Length 98;
Best Local Similarity 100.0%; Pred. No. 3e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETPTDINPAWYAGRGIRPVGR 30
    |||||
Db 23 SRAHQHSMETPTDINPAWYAGRGIRPVGR 52

RESULT 8
US-10-044-592-82
; Sequence 82, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 82
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-82

Query Match          99.4%; Score 166; DB 13; Length 98;
Best Local Similarity 100.0%; Pred. No. 3e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETPTDINPAWYAGRGIRPVGR 30
    |||||
Db 23 SRAHQHSMETPTDINPAWYAGRGIRPVGR 52

RESULT 9
US-10-044-592-84
; Sequence 84, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 84
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-84

Query Match          99.4%; Score 166; DB 13; Length 98;
Best Local Similarity 100.0%; Pred. No. 3e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETPTDINPAWYAGRGIRPVGR 30
    |||||
Db 23 SRAHQHSMETPTDINPAWYAGRGIRPVGR 52

RESULT 10
US-10-044-592-86
; Sequence 86, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 86
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-86

Query Match          99.4%; Score 166; DB 13; Length 98;
Best Local Similarity 100.0%; Pred. No. 3e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETPTDINPAWYAGRGIRPVGR 30
    |||||
Db 23 SRAHQHSMETPTDINPAWYAGRGIRPVGR 52

```

```
Db 23 SRAHQSMETRTDINPAWYAGRGIRPVGR 52
|||||
RESULT 11
US-10-044-592-88
; Sequence 88, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; SOFTWARE:
; NUMBER OF SEQ ID NOS: 96
; SEQ ID NO 88
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-88

Query Match 99.4%; Score 166; DB 13; Length 98;
Best Local Similarity 100.0%; Pred. No. 3e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQSMETRTDINPAWYAGRGIRPVGR 30
|||||
Db 23 SRAHQSMETRTDINPAWYAGRGIRPVGR 52
|||||
RESULT 12
US-10-044-592-26
; Sequence 26, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; SOFTWARE:
; NUMBER OF SEQ ID NOS: 96
; SEQ ID NO 26
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-26

Query Match 96.4%; Score 161; DB 13; Length 29;
Best Local Similarity 100.0%; Pred. No. 4.4e-16;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQSMETRTDINPAWYAGRGIRPVG 29
|||||
Db 1 SRAHQSMETRTDINPAWYAGRGIRPVG 29
|||||
RESULT 13

US-09-932-161-14
; Sequence 14, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Civealli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; TITLE OF INVENTION: Promoting Wakefulness and Sleep
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Rattus
US-09-932-161-14

Query Match 94.0%; Score 157; DB 9; Length 31;
Best Local Similarity 93.3%; Pred. No. 1.8e-15;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SRAHQSMETRTDINPAWYAGRGIRPVGR 30
|||||
Db 1 SRAHQSMETRTDINPAWYAGRGIRPVGR 30
|||||
RESULT 14
US-10-044-592-4
; Sequence 4, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; SOFTWARE:
; NUMBER OF SEQ ID NOS: 96
; SEQ ID NO 4
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Murine
US-10-044-592-4

Query Match 94.0%; Score 157; DB 13; Length 31;
Best Local Similarity 93.3%; Pred. No. 1.8e-15;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SRAHQSMETRTDINPAWYAGRGIRPVGR 30
|||||
Db 1 SRAHQSMETRTDINPAWYAGRGIRPVGR 30
|||||
RESULT 15
US-10-044-592-5
; Sequence 5, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
```

```

; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 5
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(31)
; OTHER INFORMATION: antigen
US-10-044-592-5

Query Match      94.0%; Score 157; DB 13; Length 31;
Best Local Similarity 93.3%; Pred. No. 1.8e-15;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 SRAHQHSMETRTDINPANYAGRGIRPVGR 30
      |||||  |||||  |||||  |||||  |||||
Db      1 SRAHQHSMETRTDINPANYAGRGIRPVGR 30

Search completed: August 12, 2004, 15:22:49
Job time : 54.7907 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:36 ; Search time 17.6628 Seconds
(without alignments)
90.609 Million cell updates/sec

Title: US-09-700-643A-1
Perfect score: 167
Sequence: 1 SRAHQSMIEIRTPDINPAWAGRGIRPVGR 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/prodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCTUS COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	166	99.4	31	3 US-09-105-678A-7	Sequence 7, Appli
2	166	99.4	31	3 US-09-105-678A-31	Sequence 31, Appl
3	166	99.4	31	3 US-08-776-971-5	Sequence 5, Appli
4	166	99.4	31	3 US-08-776-971-97	Sequence 97, Appl
5	166	99.4	31	3 US-09-421-208-7	Sequence 7, Appli
6	166	99.4	31	3 US-09-560-915-13	Sequence 13, Appl
7	166	99.4	31	4 US-09-105-678A-32	Sequence 32, Appl
8	166	99.4	32	3 US-08-776-971-6	Sequence 6, Appli
9	166	99.4	32	3 US-09-421-208-32	Sequence 32, Appl
10	166	99.4	32	3 US-09-105-678A-33	Sequence 33, Appl
11	166	99.4	33	3 US-08-776-971-7	Sequence 7, Appli
12	166	99.4	33	3 US-09-421-208-33	Sequence 33, Appl
13	166	99.4	33	3 US-08-776-971-1	Sequence 1, Appli
14	166	99.4	98	3 US-08-776-971-44	Sequence 44, Appl
15	166	99.4	98	3 US-08-776-971-122	Sequence 122, App
16	166	99.4	98	3 US-08-776-971-131	Sequence 131, App
17	166	99.4	98	3 US-08-776-971-136	Sequence 136, App
18	166	99.4	98	3 US-08-776-971-115	Sequence 115, App
19	162	97.0	98	3 US-08-776-971-117	Sequence 117, App
20	162	97.0	98	3 US-09-105-678A-29	Sequence 29, Appl
21	161	96.4	29	3 US-08-776-971-3	Sequence 3, Appli
22	161	96.4	29	3 US-09-421-208-29	Sequence 29, Appl
23	161	96.4	29	3 US-08-105-678A-8	Sequence 8, Appli
24	157	94.0	31	3 US-09-105-678A-37	Sequence 37, Appl
25	157	94.0	31	3 US-09-172-353-4	Sequence 4, Appli
26	157	94.0	31	3 US-08-776-971-47	Sequence 47, Appl
27	157	94.0	31	3 US-08-776-971-47	Sequence 47, Appl

28	157	94.0	31	3 US-09-421-208-8	Sequence 8, Appli
29	157	94.0	31	3 US-09-421-208-37	Sequence 37, Appl
30	157	94.0	31	4 US-09-560-915-14	Sequence 14, Appl
31	157	94.0	31	4 US-09-799-955-4	Sequence 4, Appli
32	157	94.0	32	3 US-09-105-678A-38	Sequence 38, Appl
33	157	94.0	32	3 US-08-776-971-48	Sequence 48, Appl
34	157	94.0	32	3 US-09-421-208-38	Sequence 38, Appl
35	157	94.0	33	3 US-09-105-678A-39	Sequence 39, Appl
36	157	94.0	33	3 US-08-776-971-49	Sequence 49, Appl
37	157	94.0	33	3 US-09-421-208-39	Sequence 39, Appl
38	157	94.0	83	3 US-08-776-971-45	Sequence 45, Appl
39	157	94.0	83	3 US-08-776-971-124	Sequence 124, App
40	157	94.0	83	3 US-08-776-971-137	Sequence 137, App
41	152	91.0	31	3 US-09-105-678A-9	Sequence 9, Appli
42	152	91.0	31	3 US-09-105-678A-43	Sequence 43, Appl
43	152	91.0	31	3 US-08-776-971-61	Sequence 61, Appl
44	152	91.0	31	3 US-09-421-208-9	Sequence 9, Appli
45	152	91.0	31	3 US-09-421-208-43	Sequence 43, Appl

ALIGNMENTS

RESULT 1
US-09-105-678A-7
; Sequence 7, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-105-678A-7

Query Match 99.4%; Score 166; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 9.5e-19;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQSMIEIRTPDINPAWAGRGIRPVGR 30

Db 1 SRAHQSMSEIRTPDINPAWYAGRGIRPVGR 30
|||||
RESULT 2
US-09-105-678A-31
; Sequence 31, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-105-678A-31
Query Match 99.4%; Score 166; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 9.5e-19;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQSMSEIRTPDINPAWYAGRGIRPVGR 30
Db 1 SRAHQSMSEIRTPDINPAWYAGRGIRPVGR 30
|||||
RESULT 3
US-08-776-971-5
; Sequence 5, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Habata, Yugo
; APPLICANT: Kawamata, Yuji
; APPLICANT: Hosoya, Masaki
; APPLICANT: Fujii, Ryo
; APPLICANT: Fukusumi, Shoji
; APPLICANT: Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASCTEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-08-776-971-5
Query Match 99.4%; Score 166; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 9.5e-19;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQSMSEIRTPDINPAWYAGRGIRPVGR 30
Db 1 SRAHQSMSEIRTPDINPAWYAGRGIRPVGR 30
|||||
RESULT 4
US-08-776-971-97
; Sequence 97, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Habata, Yugo
; APPLICANT: Kawamata, Yuji
; APPLICANT: Hosoya, Masaki
; APPLICANT: Fujii, Ryo
; APPLICANT: Fukusumi, Shoji
; APPLICANT: Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASCTEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-08-776-971-5
Query Match 99.4%; Score 166; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 9.5e-19;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQSMSEIRTPDINPAWYAGRGIRPVGR 30
Db 1 SRAHQSMSEIRTPDINPAWYAGRGIRPVGR 30
|||||

```

1 ;
2 ; ZIP: 02109
3 ;
4 ; COMPUTER READABLE FORM:
5 ; MEDIUM TYPE: Diskette
6 ; COMPUTER: IBM compatible
7 ; OPERATING SYSTEM: DOS
8 ; SOFTWARE: FastSEQ for Windows Version 2.0
9 ;
10 ; CURRENT APPLICATION DATA:
11 ; APPLICATION NUMBER: US/08/776,971B
12 ; FILING DATE: 06-Feb-1997
13 ; CLASSIFICATION: <Unknown>
14 ;
15 ; PRIOR APPLICATION DATA:
16 ; APPLICATION NUMBER: PCT/JP96/03821
17 ; FILING DATE: 28-DEC-1996
18 ; APPLICATION NUMBER: JP 7/343371
19 ; FILING DATE: 28-DEC-1995
20 ; APPLICATION NUMBER: JP 8/59419
21 ; FILING DATE: 15-MAR-1996
22 ; APPLICATION NUMBER: JP 8/211805
23 ; FILING DATE: 12-AUG-1996
24 ; APPLICATION NUMBER: JP 8/246573
25 ; FILING DATE: 18-SEP-1996
26 ; ATTORNEY/AGENT INFORMATION:
27 ; NAME: Conlin, David G.
28 ; REGISTRATION NUMBER: 27,026
29 ; REFERENCE/DOCKET NUMBER: 47176
30 ; TELECOMMUNICATION INFORMATION:
31 ; TELEPHONE: 617-523-3400
32 ; TELEFAX: 617-523-6440
33 ;
34 ; INFORMATION FOR SEQ ID NO: 97:
35 ; SEQUENCE CHARACTERISTICS:
36 ; LENGTH: 31 amino acids
37 ; TYPE: amino acid
38 ; STRANDEDNESS: single
39 ; TOPOLOGY: linear
40 ;
41 ; SEQUENCE DESCRIPTION: SEQ ID NO: 97:
42 ;
43 ; US-08-776-971-97
44 ;
45 ; Query Match 99.4%; Score 166; DB 3; Length 31;
46 ; Best Local Similarity 100.0%; Pred.No.9.5e-19;
47 ; Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
48 ;
49 QY 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGR 30
50 ;| | | | | | | | | | | | | | | | | | | | |
51 DB 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGR 30
52 ;| | | | | | | | | | | | | | | | | | | | |
53 ;
54 ; RESULT 5
55 ; US-09-421-208-7
56 ; Sequence 7, Application US/09421208
57 ; Patent No. 6258561
58 ; GENERAL INFORMATION:
59 ; APPLICANT: Suenaga, Masato
60 ; APPLICANT: Moriya, Takeo
61 ; APPLICANT: Tanaka, Yoko
62 ; APPLICANT: Nishimura, Osamu
63 ; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
64 ; NUMBER OF SEQUENCES: 52
65 ; CORRESPONDENCE ADDRESS:
66 ; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
67 ; STREET: 130 Water Street
68 ; CITY: Boston
69 ; STATE: MA
70 ; COUNTRY: USA
71 ; ZIP: 02109
72 ;
73 ; COMPUTER READABLE FORM:
74 ; MEDIUM TYPE: Floppy disk
75 ; COMPUTER: IBM PC compatible
76 ; OPERATING SYSTEM: PC-DOS/MS-DOS
77 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
78 ;
79 ; CURRENT APPLICATION DATA:
80 ; APPLICATION NUMBER: US/09/421,208
81 ; FILING DATE:
82 ; PRIOR APPLICATION DATA:

```

```
/ MOLECULE TYPE: peptide
/ US-09-421-208-31
/
/ Query Match 99.4%; Score 166; DB 3; Length 31;
/ Best Local Similarity 100.0%; Pred. No. 9.5e-19;
/ Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/
Qy 1 SRAHQSMWEIRTPDINPAWYAGRGIRPVGR 30
Db 1 SRAHQSMWEIRTPDINPAWYAGRGIRPVGR 30
/
/ RESULT 7
/ US-09-560-915-13
/ Sequence 13, Application US/09560915
/ Patent No. 6383764
/ GENERAL INFORMATION:
/ APPLICANT: Lin, Steven
/ TITLE OF INVENTION: Therapeutic Compositions and Methods
/ FILE REFERENCE: P-UC 3534
/ CURRENT APPLICATION NUMBER: US/09/560,915
/ CURRENT FILING DATE: 2000-04-28
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 13
/ LENGTH: 31
/ TYPE: PRT
/ ORGANISM: Bos taurus
/ US-09-560-915-13
/
/ Query Match 99.4%; Score 166; DB 4; Length 31;
/ Best Local Similarity 100.0%; Pred. No. 9.5e-19;
/ Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/
Qy 1 SRAHQSMWEIRTPDINPAWYAGRGIRPVGR 30
Db 1 SRAHQSMWEIRTPDINPAWYAGRGIRPVGR 30
/
/ RESULT 8
/ US-09-105-678A-32
/ Sequence 32, Application US/09105678A
/ Patent No. 6103882
/ GENERAL INFORMATION:
/ APPLICANT: Suenaga, Masato
/ APPLICANT: Moriya, Takeo
/ APPLICANT: Tanaka, Yoko
/ APPLICANT: Nishimura, Osamu
/ TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
/ NUMBER OF SEQUENCES: 52
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
/ STREET: 130 Water Street
/ CITY: Boston
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02109
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/105,678A
/ FILING DATE: 26-JUN-1998
/ PRIORITY APPLICATION DATA:
/ APPLICATION NUMBER: JP 172118/1997
/ FILING DATE: 27-JUN-1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Conlin, David G.
/ REGISTRATION NUMBER: 27,026
/
/ Query Match 99.4%; Score 166; DB 3; Length 32;
/ Best Local Similarity 100.0%; Pred. No. 9.9e-19;
/ Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/
Qy 1 SRAHQSMWEIRTPDINPAWYAGRGIRPVGR 30
Db 1 SRAHQSMWEIRTPDINPAWYAGRGIRPVGR 30
/
/ RESULT 9
/ US-08-776-971-6
/ Sequence 6, Application US/08776971B
/ Patent No. 6228984
/ GENERAL INFORMATION:
/ APPLICANT: Hinuma, Shuji
/ APPLICANT: Habata, Yugo
/ APPLICANT: Kawamata, Yuji
/ APPLICANT: Hosoya, Masaki
/ APPLICANT: Fujii, Ryo
/ APPLICANT: Fukusumi, Shoji
/ APPLICANT: Kitada, Chieko
/ TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
/ NUMBER OF SEQUENCES: 140
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
/ STREET: 130 Water Street
/ CITY: Boston
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02109
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/776,971B
/ FILING DATE: 06-Feb-1997
/ CLASSIFICATION: <Unknown>
/ PRIORITY APPLICATION DATA:
/ APPLICATION NUMBER: PCT/JF96/03821
/ FILING DATE: 28-DEC-1996
/ APPLICATION NUMBER: JP 7/343371
/ FILING DATE: 28-DEC-1995
/ APPLICATION NUMBER: JP 8/59419
/ FILING DATE: 15-MAR-1996
/ APPLICATION NUMBER: JP 8/211805
/ FILING DATE: 12-AUG-1996
/ APPLICATION NUMBER: JP 8/246573
/ FILING DATE: 18-SEP-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Conlin, David G.
/ REGISTRATION NUMBER: 27,026
/ REFERENCE/DOCKET NUMBER: 47176
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-523-3400
/ TELEFAX: 617-523-6440
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 32 amino acids
```



```

;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-08-776-971-6
Query Match          99.4%; Score 166; DB 3; Length 32;
Best Local Similarity 100.0%; Pred. No. 9.9e-19;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQSHMEIRTPDINPAWAGRGIRPVGR 30
DB 1 SRAHQSHMEIRTPDINPAWAGRGIRPVGR 30

RESULT 10
US-09-421-208-32
; Sequence 32, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 27,026
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 48466-342
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-421-208-32
Query Match          99.4%; Score 166; DB 3; Length 32;
Best Local Similarity 100.0%; Pred. No. 9.9e-19;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQSHMEIRTPDINPAWAGRGIRPVGR 30
DB 1 SRAHQSHMEIRTPDINPAWAGRGIRPVGR 30

RESULT 11
US-09-105-678A-33
; Sequence 33, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-105-678A-33
Query Match          99.4%; Score 166; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQSHMEIRTPDINPAWAGRGIRPVGR 30
DB 1 SRAHQSHMEIRTPDINPAWAGRGIRPVGR 30

RESULT 12
US-08-776-971-7
; Sequence 7, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Habata, Yugo
; APPLICANT: Kawamata, Yuji
; APPLICANT: Hosoya, Masaki
; APPLICANT: Fujii, Ryo
; APPLICANT: Fukusumi, Shoji
; APPLICANT: Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 27,026
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 48466-342
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-421-208-32
Query Match          99.4%; Score 166; DB 3; Length 32;
Best Local Similarity 100.0%; Pred. No. 9.9e-19;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQSHMEIRTPDINPAWAGRGIRPVGR 30
DB 1 SRAHQSHMEIRTPDINPAWAGRGIRPVGR 30
```

```
;
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
;
; US-08-776-971-7
;
; Query Match 99.4%; Score 166; DB 3; Length 33;
; Best Local Similarity 100.0%; Pred. NO. 1e-18;
; Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 SRAHQSHWEIRTPDINPAWYAGRGIRPVGR 30
; Db 1 SRAHQSHWEIRTPDINPAWYAGRGIRPVGR 30
;
; RESULT 13
; US-09-421-208-33
; Sequence 33, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
```

```
;
; APPLICATION NUMBER: US/09/421,208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
; US-09-421-208-33
;
; Query Match 99.4%; Score 166; DB 3; Length 33;
; Best Local Similarity 100.0%; Pred. NO. 1e-18;
; Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 SRAHQSHWEIRTPDINPAWYAGRGIRPVGR 30
; Db 1 SRAHQSHWEIRTPDINPAWYAGRGIRPVGR 30
;
; RESULT 14
; US-08-776-971-1
; Sequence 1, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; Habata, Yugo
; Kawamata, Yuji
; Hosoya, Masaki
; Fujii, Ryo
; Fukusumi, Shoji
; Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:36 ; Search time 13.6977 Seconds
(without alignments)
217.697 Million cell updates/sec

Title: US-09-700-643a-2
Perfect score: 166
Sequence: 1 SRTHSHMEIRTPDINPAWYASRGIRPVGR 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:***
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	143	86.1	83	JC7607	prolactin-releasin
2	56.5	34.0	684	F83376	conserved hypothet
3	53.5	32.2	314	B70569	hypothetical prote
4	53	31.9	637	D87559	sensory box histid
5	53	31.9	790	T47959	hypothetical prote
6	52	31.3	785	F69099	sensory transducti
7	50	30.1	123	S77900	hypothetical prote
8	50	30.1	128	S76955	hypothetical prote
9	49.5	29.8	176	S67150	hypothetical prote
10	49	29.5	72	E91002	probable regulator
11	48.5	29.2	303	AH2016	hypothetical prote
12	48	28.9	118	AC3169	hypothetical prote
13	48	28.9	220	C83292	probable glutathio
14	48	28.9	335	S70671	lipopolysaccharide
15	48	28.9	348	T21648	hypothetical prote
16	48	28.9	455	D70885	hypothetical prote
17	48	28.9	1084	T33759	probable aldc prot
18	47.5	28.6	345	D84012	hypothetical prote
19	47.5	28.6	393	AB2564	N-acetylglutamate
20	47.5	28.6	401	A97446	conserved hypothet
21	47.5	28.6	503	A82193	hypothetical prote
22	47.5	28.6	533	A33111	sun/nucleolar prot
23	47	28.3	159	G82669	segmentation prote
24	47	28.3	215	B87577	ubiquitome biosynt
25	47	28.3	284	F71015	glutathione S-tran
26	47	28.3	333	H82852	hypothetical prote
27	47	28.3	501	T48336	hydroxybenzoate oc
28	47	28.3	938	C84480	hypothetical prote
29	47	28.3	4589	T14314	dynein beta heavy

30 46.5 28.0 240 2 D64688 probable 1-acylgly
31 46.5 28.0 779 2 T49717 related to BCS1 pr
32 46.5 28.0 957 2 A84089 hypothetical prote
33 46.5 28.0 1495 2 T31434 densin-180 - rat
34 46.5 28.0 1607 2 T13350 hypothetical prote
35 46.5 28.0 1693 1 MNWHE genome polyprotein
36 46.5 28.0 4957 2 T03455 ALR protein - huma
37 46.5 28.0 5262 2 T03454 ALR protein - huma
38 46 27.7 256 2 F70812 probable lpgR prot
39 46 27.7 342 2 B64395 malic acid transpo
40 46 27.7 347 2 H64371 malic acid transpo
41 46 27.7 419 2 AH3166 hypothetical prote
42 46 27.7 688 2 AI2516 hypothetical prote
43 46 27.7 698 2 T39050 hypothetical prote
44 46 27.7 1004 2 H87112 glutamate-ammonia-
45 46 27.7 1236 2 T50904 Mg protoporphyzin

ALIGNMENTS

RESULT 1

JC7607
prolactin-releasing peptide - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C:Accession: JC7607
Riyamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Hosoya, T.;
Biochem. Biophys. Res. Commun. 281, 53-56, 2001
A>Title: Isolation and characterization of the rat prolactin-releasing peptide gene: Mu
A:Reference number: JC7607; MUID:21092785; PMID:11178959
A:Contents: Spleen
A:Accession: JC7607
A:Molecule type: DNA
A:Residues: 1-83 <YAM>
A:Cross-references: DDBJ:AB040612; DDBJ:AB040613
C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior pi
release, and stimulation of ACTH secretion from the pituitary.
C:Genetics:
A:Gene: PrRP
A:introns: 33/1

Query Match 86.1%; Score 143; DB 2; Length 83;
Best Local Similarity 83.3%; Pred. No. 5e-14;
Matches 25; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 SRTHSHMEIRTPDINPAWYASRGIRPVGR 30
|||:|||||
DB 22 SRAHQSHMETPTDINPAWYTGIRPVGR 51
|||:|||||

RESULT 2

F83376
conserved hypothetical protein PA2151 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83376
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-984, 2000
A>Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10384043
A:Accession: F83376
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-664 <STO>
A:Cross-references: GB:AE004642; GB:AE004091; NID:g9948163; PIDN:AG05539.1; GSPDB:GN001
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: PA2151

Query Match 34.0%; Score 56.5; DB 2; Length 664;

```

Best Local Similarity 45.8%; Pred. No. 3.1;
Matches 11; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 4 HRRSMEIRTPDINPAWYASRGIRP 27
Db 470 YRPNFFVTPDINP-WFLQSRGP 492

RESULT 3
B70569
hypothetical protein RV3485c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: B70569
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Seares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: B70569
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-314 <COL>
A:Cross-references: GB:Z95390; GB:AL123456; NID:G3261766; PIDN:CAB08708.1; PID:G2104408
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV3485c
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
F:46-227/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 32.2%; Score 53.5; DB 2; Length 314;
Best Local Similarity 27.5%; Pred. No. 3.8;
Matches 11; Conservative 5; Mismatches 11; Indels 13; Gaps 1;

QY 1 SRTHR-----HSMEIRTPDINPAWYASRGIRP 27
Db 187 SNTHRFVGAIVTKSAVDHMKLADELGFSWRVNSIRP 226

RESULT 4
D87559
sensory box histidine kinase/response regulator [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: D87559
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: AB7249; MUID:21173698; PMID:11259647
A:Accession: D87559
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-637 <STO>
A:Cross-references: GB:AE005673; NID:GI3424056; PIDN:AAK24472.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC2501

Query Match 31.9%; Score 53; DB 2; Length 637;
Best Local Similarity 48.0%; Pred. No. 9.7;
Matches 12; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 4 HRRSMEIRTPDINPAWYASRGIRP 28
Db 22 HRDSDLRSPAINPAIRVRLRAV 46

RESULT 5
T47959
hypothetical protein F15G16.60 - Arabidopsis thaliana

```

```

C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47959
R:De Haan, M.; Maarse, A.C.; Grivell, L.A.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quet, submitted to the Protein Sequence Database, January 2000
A:Reference number: Z24480
A:Accession: T47959
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-790 <DEH>
A:Cross-references: EMBL:AL132959
A:Experimental source: cultivar Columbia; BAC clone F15G16
C:Genetics:
A:Map position: 3
A:Introns: 39/1; 678/2; 698/3; 773/2
A:Note: F15G16.60

Query Match 31.9%; Score 53; DB 2; Length 790;
Best Local Similarity 55.6%; Pred. No. 12;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 13 PDINPAWYASRGIRPVGR 30
Db 366 PPHNPTYGSRGLQPHGR 383

RESULT 6
F69099
sensory transduction histidine kinase - Methanobacterium thermoautotrophicum (strain Del)
C:Species: Methanobacterium thermoautotrophicum
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: F69099
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, N.; Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: function.
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: F69099
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-785 <MTH>
A:Cross-references: GB:AE000805; GB:AE000666; NID:G2621213; PIDN:AAB84680.1; PID:G2621213
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH174

Query Match 31.3%; Score 52; DB 2; Length 785;
Best Local Similarity 46.4%; Pred. No. 17;
Matches 13; Conservative 1; Mismatches 12; Indels 2; Gaps 1;

QY 4 HRRSMEIRTPDINPAW--YASRGIRPVG 29
Db 412 HRATFIRRPDGNRVVVEYVDPIRTDG 439

RESULT 7
S77900
hypothetical protein 1 - Clostridium barkeri (fragment)
C:Species: Clostridium barkeri
C>Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 05-Nov-1999
C:Accession: S77900; S43551
R:Beatrice, B.; Zeider, O.; Linder, D.; Buckel, W.
Eur. J. Biochem. 221, 101-109, 1994
A:Title: Cloning, sequencing and expression of the gene encoding the coenzyme B(12)-dependent
A:Reference number: S43237; MUID:94222050; PMID:8168499
A:Accession: S77900
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-123 <BEA>
A:Cross-references: EMBL:X77484; NID:G472895; PIDN:CAA54624.1; PID:G472896

Query Match 30.1%; Score 50; DB 2; Length 123;

```

gaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: E91002
A>Status: preliminary
A;Molecule type: DNA
A;Residues: 1-72 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB36412.1; PID:gl3362458; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RMD 0509952
C;Genetics:
A;Gene: ECs2989

Query Match 29.5%; Score 49; DB 2; Length 72;
Best Local Similarity 34.5%; Pred. No. 3.5;
Matches 10; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

QY 1 SRTHRSMEIRTPDINPWAYASRGIRPVG 29
||| : ||||| :
DB 27 SRIANYELNIPTGLNDCRMIVEGLRKLK 55
||| : ||||| :

RESULT 11
AH2016
hypothetical protein all1686 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 24-Nov-2003
C;Accession: AH2016
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH2016
A>Status: preliminary
A;Molecule type: DNA
A;Residues: 1-303 <KUR>
A;Cross-references: GB:BA0000119; PIDN:BAB78052.1; PID:gi7135506; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all1686
C;Superfamily: tartrate-resistant acid phosphatase

Query Match 29.2%; Score 48.5; DB 2; Length 303;
Best Local Similarity 48.4%; Pred. No. 20;
Matches 15; Conservative 1; Mismatches 10; Indels 5; Gaps 2;

QY 4 HRHSMEIRTPDINPAWY-----ASRGIRPVGR 30
||||| : |
DB 226 HEHSYS-RTRAIIDGTTLTCGAGAGNRVPVGR 255
||||| : |

RESULT 12
AC3169
hypothetical protein Atu5077 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AC3169
R;Wood, D.W.; Serubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AC3169
A>Status: preliminary
A;Molecule type: DNA
A;Residues: 1-118 <KUR>
A;Cross-references: GB:AE008687; PIDN:AAL45769.1; PID:gi7743503; GSPDB:GN00188
A;Experimental source: strain C58 (Dupont)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 8.11047 Seconds
(without alignments)
199.024 Million cell updates/sec

Title: US-09-700-643A-2

Perfect score: 166
Sequence: 1 SRTHRSMETPTDINPAWYASRGIRPVGRX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	165	99.4	87	1 PRRP_HUMAN	P81277 homo sapien
2	152	91.6	98	1 PRRP_BOVIN	P81264 bos taurus
3	143	86.1	83	1 PRRP_RAT	P81278 rattus norv
4	54	32.5	428	1 NER3_HUMAN	Q9UC49 homo sapien
5	52	31.3	288	1 Y587_PASMU	Q9CM56 pasteurella
6	50	30.1	428	1 NER3_BOVIN	Q97859 bos-taurus
7	47.5	28.6	345	1 ARG3_BACHD	Q9K8V2 bacillus ha
8	47.5	28.6	533	1 CNC_DROME	P20482 drosophila
9	47	28.3	402	1 EX7L_STRCO	Q9FBM3 streptomyce
10	46.5	28.0	240	1 PLSC_HELPY	O25903 helicobacte
11	46.5	28.0	333	1 SIR4_MOUSE	Q8R216 mus musculu
12	46.5	28.0	535	1 PYRG_BACAA	Q81JW1 bacillus an
13	46.5	28.0	1693	1 POLN_HEVBU	P29324 hepatitis e
14	46.5	28.0	1693	1 POLN_HEVNY	Q04610 hepatitis e
15	46.5	28.0	5262	1 MLL2_HUMAN	O14686 homo sapien
16	46	27.7	342	1 Y762_METJA	Q58172 methanococc
17	46	27.7	346	1 ARG3_RHOBA	Q7UV14 rhodospirell
18	46	27.7	347	1 Y576_METJA	Q57996 methanococc
19	46	27.7	383	1 CYCR_CHRVI	O82947 chromatiu
20	46	27.7	413	1 EX7L_COREF	Q8FGP1 corynebacte
21	46	27.7	417	1 EX7L_CORGL	O8NM33 corynebacte
22	46	27.7	1400	1 RIF1_SCHPO	Q96UP3 schizosacch
23	45.5	27.4	239	1 6PGL_SVNY3	P74618 synchocyst
24	45.5	27.4	407	1 Y116_WCTUT	P72052 mycobacteri
25	45	27.1	213	1 SRN3_YEAST	Q99176 saccharomyc
26	45	27.1	321	1 HTPX_AGRTS	Q8UBMS agrobacteri
27	45	27.1	350	1 HMRO_DROME	P10181 drosophila
28	45	27.1	510	1 YCGE_ECOLI	P29013 escherichia
29	45	27.1	798	1 UNR_RAT	P18395 rattus norv
30	44.5	26.8	137	1 YMAF_BACSU	O31794 bacillus su
31	44.5	26.8	850	1 YG12_BPB03	Q37893 bacterioph
32	44.5	26.8	1485	1 YH85_SCHPO	Q3PSN0 schizosacch
33	44.5	26.8	1882	1 Y468_MYCPN	P75109 mycoplasma

34 44 26.5 137 1 SMR2_RAT P18897 rattus norv
35 44 26.5 364 1 YAIM_ECOLI P77562 escherichia
36 44 26.5 386 1 CRTY_AGRAU P54974 agrobacteri
37 44 26.5 476 1 YAAJ_ECOLI P30143 escherichia
38 44 26.5 581 1 POL_MLVRK P31795 radiation m
39 44 26.5 591 1 PYRG_HUMAN P17812 homo sapien
40 44 26.5 719 1 NRPI_YEAST P03357 akr murine
41 44 26.5 843 1 POL_MLVAK P32770 saccharomyc
42 44 26.5 1087 1 XPO7_HUMAN Q9UIA9 homo sapien
43 44 26.5 1087 1 XPO7_MOUSE Q9EPK7 mus musculu
44 44 26.5 1196 1 POL_MLVAV P03356 akr murine
45 44 26.5 1196 1 POL_MLVRO P12227 radiation m

ALIGNMENTS

RESULT 1

PRRP_HUMAN
ID PRRP_HUMAN STANDARD; PRT; 87 AA.
AC P81277;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-releasing peptide PrRP20].
DE releasing peptide PrRP20].
GN PRH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98268781; PubMed=9607765;
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;
RT "A prolactin-releasing peptide in the brain.";
RL Nature 393:272-276(1998).
RN [2]
RP TISSUE SPECIFICITY.
RX MEDLINE=99426652; PubMed=10498338;
RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S., Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H., Sumino Y., Fujino M.;
RT "Tissue distribution of prolactin-releasing peptide (PrRP) and its receptor.";
RL Regul. Pept. 83:1-10(1999).
CC -!- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.
CC -!- TISSUE SPECIFICITY: Medulla oblongata and hypothalamus.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB015419; BAA29027.1; -.
DR MIM; 602663; -.
DR GO; GO:0005180; P:peptide hormone; TAS.
DR Hormone; Amication; Signal.
FT SIGNAL 1 22 BY SIMILARITY.
FT PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRRP31.
FT PROPEP 58 87 PROLACTIN-RELEASING PEPTIDE PRRP20.
FT MOD_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).
SQ SEQUENCE 87 AA; 9639 MW; 229A2F3F50CF981B CRC64;

```

Query Match          99.4%; Score 165; DB 1; Length 87;
Best Local Similarity 100.0%; Pred. No. 5.5e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHSHMEIRTPDINPAVYASGIRPVGR 30
    |||
Db 23 SRTHSHMEIRTPDINPAVYASGIRPVGR 52

RESULT 2
PRRP BOVIN
ID _PRRP_BOVIN STANDARD; PRT; 98 AA.
AC P81264;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing
DE hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-
DE releasing peptide PrRP20].
GN PRH.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
CX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.
RC TISSUE=Brain;
RX MEDLINE=98268781; PubMed=9607765;
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
RA "A prolactin-releasing peptide in the brain.";
RL Nature 393:272-276(1998).
CC -!- FUNCTION: Stimulates prolactin (PRL) release and regulates the
CC expression of prolactin through its receptor GPR10. May stimulate
CC lactotrophs directly to secrete PRL.
CC -!- TISSUE SPECIFICITY: Medulla oblongata and hypothalamus.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB015417; BAA29025.1; -
KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.
FT SIGNAL 1 22
FT PEPTIDE 23 53 Prolactin-releasing peptide PrRP31.
FT PROPEP 58 98 Prolactin-releasing peptide PrRP20.
FT MOD_RES 53 53 Amidation (G-54 provide amide group).
FT VARSPLIC 98 AA; 10544 MW; 08AC35A13B0FA908 CRC64;
SQ SEQUENCE 98 AA; 10544 MW; 08AC35A13B0FA908 CRC64;

Query Match          91.6%; Score 152; DB 1; Length 98;
Best Local Similarity 90.0%; Pred. No. 5.7e-16;
Matches 27; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRTHSHMEIRTPDINPAVYASGIRPVGR 30
    |||
Db 23 SRTHSHMEIRTPDINPAVYASGIRPVGR 52

RESULT 3
PRRP RAT
ID _PRRP_RAT STANDARD; PRT; 83 AA.
AC P81278; O8K3Y0.
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing
DE hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-
DE releasing peptide PrRP20].
GN PRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=98268781; PubMed=9607765;
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
RA "A prolactin-releasing peptide in the brain.";
RL Nature 393:272-276(1998).
CC -!- FUNCTION: Stimulates prolactin (PRL) release and regulates the
CC expression of prolactin through its receptor GPR10. May stimulate
CC lactotrophs directly to secrete PRL.
CC -!- TISSUE SPECIFICITY: Medulla oblongata and hypothalamus.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB015418; BAA29026.1; -
KW Hormone; Amidation; Signal; Cleavage on pair of basic residues;
FT SIGNAL 1 21
FT PEPTIDE 22 52 Prolactin-releasing peptide PrRP31.
FT PROPEP 57 83 Prolactin-releasing peptide PrRP20.
FT MOD_RES 52 52 Amidation (G-53 provide amide group).
FT VARSPLIC 53 83 TPDLNPAWYTGIRPVGRPRRRATPRDTVIGLQSCULPL
FT DGETKFSQRC -> SECLTYGKQPLTSFPHFTSQMPP (in
FT isoform 2).
FT /FTID=VSP 004370.
FT /DOC75A264EEB4F29 CRC64;
SQ SEQUENCE 83 AA; 9215 MW; 08C75A264EEB4F29 CRC64;

Query Match          86.1%; Score 143; DB 1; Length 83;
Best Local Similarity 83.3%; Pred. No. 1.1e-14;
Matches 25; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing
DE hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-
DE releasing peptide PrRP20].
GN PRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=98268781; PubMed=9607765;
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
RA "A prolactin-releasing peptide in the brain.";
RL Nature 393:272-276(1998).
CC -!- FUNCTION: Stimulates prolactin (PRL) release and regulates the
CC expression of prolactin through its receptor GPR10. May stimulate
CC lactotrophs directly to secrete PRL.
CC -!- TISSUE SPECIFICITY: Medulla oblongata and hypothalamus.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB015418; BAA29026.1; -
KW Hormone; Amidation; Signal; Cleavage on pair of basic residues;
FT SIGNAL 1 21
FT PEPTIDE 22 52 Prolactin-releasing peptide PrRP31.
FT PROPEP 57 83 Prolactin-releasing peptide PrRP20.
FT MOD_RES 52 52 Amidation (G-53 provide amide group).
FT VARSPLIC 53 83 TPDLNPAWYTGIRPVGRPRRRATPRDTVIGLQSCULPL
FT DGETKFSQRC -> SECLTYGKQPLTSFPHFTSQMPP (in
FT isoform 2).
FT /FTID=VSP 004370.
FT /DOC75A264EEB4F29 CRC64;
SQ SEQUENCE 83 AA; 9215 MW; 08C75A264EEB4F29 CRC64;

Query Match          86.1%; Score 143; DB 1; Length 83;
Best Local Similarity 83.3%; Pred. No. 1.1e-14;
Matches 25; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 1 SRTHRSMETPTDINPAWYASRGIRPV 30
DB 22 SRAHQSMETPTDINPAWYTGIRPVGR 51

RESULT 4
NER3_HUMAN
ID NER3_HUMAN STANDARD; PRT; 428 AA.
AC G9UQ49; C9NQEL;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sialidase 3 (EC 3.2.1.18) (Membrane sialidase)
DE (N-acetyl-alpha-neuraminidase 3).
GN NEU3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RT TISSUE=Brain;
RX MEDLINE=9933533; PubMed=10405317;
RA Wada T., Yoshikawa Y., Tokuyama S., Kuwabara M., Akita H., Miyagi T.;
RT "Cloning, expression, and chromosomal mapping of a human ganglioside
RL sialidase."
RL Biochem. Biophys. Res. Commun. 261:21-27(1999).
RN [2]
RP SEQUENCE FROM N.A.
RT TISSUE=Skeletal muscle;
RX PubMed=10861246;
RA Monti E., Bassi M.T., Papini N., Riboni M., Manzoni M., Venerando B.,
RA Crcchi G., Preti A., Ballabio A., Tettamanti G., Borsani G.;
RT "Identification and expression of NEU3, a novel human sialidase
RT associated to the plasma membrane."
RL Biochem. J. 349:343-351(2000).
CC -I- FUNCTION: Plays a role in modulating the ganglioside content of
CC the lipid bilayer at the level of membrane-bound sialyl
CC glycoconjugates.
CC -I- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2-3)-, alpha-(2-6)-,
CC alpha-(2-8)-glycosidic linkages of terminal sialic residues in
CC oligosaccharides, glycoproteins, glycolipids, colominic acid and
CC synthetic substrates.
CC -I- SUBCELLULAR LOCATION: Membrane-associated.
CC -I- TISSUE SPECIFICITY: Highly expressed in skeletal muscle, testis,
CC adrenal gland and thymus, followed by pancreas, liver, heart and
CC thymus. Weakly expressed in kidney, placenta, brain and lung.
CC -I- MISCELLANEOUS: Optimum pH is 3.8.
CC -I- SIMILARITY: Belongs to family 33 of glycosyl hydrolases.
CC -I- SIMILARITY: Contains 3 ENR repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB008185; BAA82611.1; -.
CC ENBL; Y18563; CAB96131.1; ALT_INIT.
CC Genew; HGNC:7760; NEU3.
CC MIM; 604617; -.
CC GO; GO:0005887; C: integral to plasma membrane; TAS.
CC GO; GO:0003824; F: catalytic activity; TAS.
CC GO; GO:0006689; P: ganglioside catabolism; TAS.
CC InterPro; IPR002860; GH_ENR.
CC Pfam; PF02012; ENR; 3.
CC Hydrolase; Glycosidase; Membrane; Repeat.
KW REPEAT 129 140 ENR 1.
FT REPEAT 203 214 ENR 2.
FT REPEAT 254 265 ENR 3.

QY 2 RTHRSMETPTDINPAWYASRGIRPV 28
DB 195 KTRPHSLMYSDDLGVTWHRGLRPM 221

Query Match 32.5%; Score 54; DB 1; Length 428;
Best Local Similarity 37.0%; Pred. No. 1.7;
Matches 10; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

RESULT 5
Y587_PASMU
ID Y587_PASMU STANDARD; PRT; 288 AA.
AC Q9CN56;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein PM0587.
GN PM0587.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RT STRAIN=PM70;
RC MEDLINE=21145866; PubMed=11248100;
RX May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida PM70."
CC -I- SIMILARITY: Belongs to the fructosamine kinase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE006094; AAK02671.1; -.
CC InterPro; IPR005581; Fructosamin_kin.
CC Pfam; PF03881; Fructosamin_kin; 1.
CC Hypothetical protein; Transferase; Kinase; Complete proteome.
KW SEQUENCE 288 AA; 33778 MW; F4D2F6C26014D940 CRC64;
CC
Query Match 31.3%; Score 52; DB 1; Length 288;
Best Local Similarity 37.5%; Pred. No. 2.2;
Matches 9; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 5 RTHSMETPTDINPAWYASRGIRPV 28
DB 20 KHEKIHGEMHEWIDDGIOPV 43

RESULT 6
NER3_BOVIN
ID NER3_BOVIN STANDARD; PRT; 428 AA.
AC Q97859.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

DE Sialidase 3 (EC 3.2.1.18) (Membrane sialidase) (Ganglioside sialidase)
 DE (N-acetyl-alpha-neuraminidase 3).
 GN NEU3.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.
 RC TISSUE=Brain;
 RX MEDLINE=99143165; PubMed=9988745;
 RA Miyagi T., Wada T., Iwamatsu A., Hata K., Yoshikawa Y., Tokuyama S.,
 RA Sawada M.;
 RT Molecular cloning and characterization of a plasma membrane-
 RT associated sialidase specific for gangliosides.";
 RL J. Biol. Chem. 274:5004-5011(1999).
 CC -!- FUNCTION: Plays a role in modulating the ganglioside content of
 CC the lipid bilayer at the level of membrane-bound sialyl
 CC glycoconjugates.
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2-3)-, alpha-(2-6)-,
 CC alpha-(2-8)-glycosidic linkages of terminal sialic residues in
 CC oligosaccharides, glycoproteins, glycolipids, colominic acid and
 CC synthetic substrates.
 CC -!- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
 CC -!- TISSUE SPECIFICITY: Expressed in brain.
 CC -!- SIMILARITY: Belongs to family 33 of glycosyl hydrolases.
 CC -!- SIMILARITY: Contains 3 BNR repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB008184; BAA75071.1; -
 DR InterPro; IPR002860; GH_BNR.
 DR Pfam; PF02012; BNR. 3.
 KW Hydrolase; Glycosidase; Membrane; Repeat.
 FT REPEAT 129 140 BNR 1.
 FT REPEAT 203 214 BNR 2.
 FT REPEAT 254 265 BNR 3.
 FT SITE 24 27 FRIP MOTIF.
 FT ACT_SITE 25 25 By similarity.
 FT ACT_SITE 45 45 Potential.
 FT ACT_SITE 50 50 Potential.
 FT ACT_SITE 87 87 Potential.
 FT ACT_SITE 225 225 Potential.
 FT ACT_SITE 245 245 Potential.
 FT ACT_SITE 341 341 By similarity.
 FT ACT_SITE 371 371 Potential.
 FT ACT_SITE 388 388 Potential.
 SQ SEQUENCE 428 AA; 47916 MW; 418B34F3245A8F21 CRC64;
 Query Match 30.1%; Score 50; DB 1; Length 428;
 Best Local Similarity 33.3%; Pred. No. 6.8;
 Matches 9; Conservative 6; Mismatches 12; Indels 0; Gaps 0;
 QY 2 RTHRSMETPTDINPAWYASRGIRPV 28
 Db 195 RARPHSLMIVSDDLGGATWHGRLIKPM 221
 RESULT 7
 ARGC_BACHD STANDARD; PRT; 345 AA.
 AC Q9K8V2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) (AGPR) (N-

DE acetyl-glutamate semialdehyde dehydrogenase) (NAGSA dehydrogenase).
 GN ARGC OR BH2900.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kunara S.,
 RA Horikoshi K.;
 RT Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 CC -!- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(+) + phosphate = N-acetyl-S-glutamyl phosphate + NADPH.
 CC -!- PATHWAY: Arginine biosynthesis; third step.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -!- SIMILARITY: Belongs to the NAGSA dehydrogenase family. Subfamily 1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AP001517; BAB06619.1; -
 DR PIR; D84012; D84012.
 DR HAMAP; MF 00150; -; 1.
 DR InterPro; IPR000706; AGPR_act site.
 DR InterPro; IPR000534; Semialdh_dh.
 DR Pfam; PF01118; Semialdh_dh.1.
 DR Pfam; PF02774; Semialdh_dhC.1.
 DR PRODOM; PD003765; AGPR_act_site.1.
 DR PROSITE; PS01224; AGRC; 1.
 DR Arginine biosynthesis; Oxidoreductase; NADP; Complete proteome.
 FT ACT_SITE 149 149 BY SIMILARITY.
 SQ SEQUENCE 345 AA; 38188 MW; 3E9F45DD09FC68EA CRC64;
 Query Match 28.8%; Score 47.5; DB 1; Length 345;
 Best Local Similarity 45.5%; Pred. No. 13;
 Matches 10; Conservative 2; Mismatches 9; Indels 1; Gaps 1;
 QY 8 MEIRTPDINPAWYASRGIRPVG 29
 Db 101 LRINEPDVVEAWY-KRQAPVVG 121
 RESULT 8
 CNC.DROME STANDARD; PRT; 533 AA.
 ID CNC.DROME
 AC P20182;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Segmentation protein cap'n/collar.
 GN CNC.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92001535; PubMed=1911393;
 RA Mohler J., Vani K., Leung S., Epstein A.;
 RT "Segmentally restricted, cephalic expression of a leucine zipper gene
 RT during Drosophila embryogenesis.";
 RL Mech. Dev. 34:3-9(1991).

```

CC -!- FUNCTION: Plays a role in cephalic patterning. Probable subunit
CC of a heterodimeric regulatory protein involved in the control of
CC head morphogenesis.
CC -!- DEVELOPMENTAL STAGE: Localized to the mandibular segment and the
CC hypopharyngeal and labral primordia first detectable in late
CC blastoderm stages.
CC -!- SIMILARITY: Belongs to the bZIP family. CNC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M37495; AAB59246.1; -
CC HSP; P34707; LSKN
CC TRANSFAC; T01998; -
CC FlyBase; FBgn0000338; cnc.
CC GO; GO:0005634; C:nucleus; IDA.
CC GO; GO:0003677; F:DNA binding; IMP.
CC GO; GO:0007350; P:blastoderm segmentation; IMP.
CC GO; GO:0007310; P:ooocyte dorsal/ventral axis determination; IMP.
CC GO; GO:0008103; P:ooocyte microtubule cytoskeleton polarization; IMP.
CC GO; GO:0042070; P:ooocyte nucleus anchoring; IMP.
CC GO; GO:0008359; P:regulation of bicoid mRNA localization; IMP.
CC GO; GO:0007317; P:regulation of pole plasm mRNA localization. .; IMP.
CC InterPro; IPR008917; Euk transcr_DNA.
CC InterPro; IPR004827; TF_bZIP.
CC Pfam; PF00170; bZIP; 1.
CC SMART; SM00338; BSLZ; 1.
CC PROSITE; PS00217; bZIP; 1.
CC PROSITE; PS00036; bZIP_BASIC; 1.
CC Transcription regulation; Activator; DNA-binding; Nuclear protein;
CC Developmental protein.
CC FT DNA BIND 351 369 BASIC MOTIF.
CC FT DOMAIN 387 408 LEUCINE-ZIPPER.
CC FT DOMAIN 497 529 GLN-RICH.
CC FT CONFLICT 349 349 I -> L (IN REF. 1; AAB59246).
CC SEQUENCE 533 AA; 56948 MW; EADFF9A5D6CA5C5F CRC64;
CC -----
Query Match 28.6%; Score 47.5; DB 1; Length 533;
Best Local Similarity 37.5%; Pred. No. 21;
Matches 9; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 4 HRRHMB-IRTPDINPAWYASRGIR 26
||| : : ||| : :
DB 409 HRRVFQYLRDPEGNPCWADYSLQ 432

RESULT 9
EX1L_STRCO STANDARD; PRT; 402 AA.
AC Q9FBN3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
DE (Exonuclease VII large subunit).
GN XSEA OR SC05056 OR SC7.29C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OC NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,

```

```

RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Reil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -!- FUNCTION: Bidirectionally degrades single-stranded DNA into large
CC acid-insoluble oligonucleotides, which are then degraded further
CC into small acid-soluble oligonucleotides (By similarity).
CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5' to 3'-
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.
CC -!- SUBUNIT: Heterooligomer composed of large and small subunits (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the xseA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL939122; CAC05901.1; -
CC HAMAP; MF_00378; -; 1.
CC InterPro; IPR003753; Exonuc_VII_L.
CC InterPro; IPR004365; tRNA anti.
CC Pfam; PF02601; Exonuc_VII_L; 1.
CC Pfam; PF01336; tRNA_anti; 1.
CC TIGRFAMs; TIGR00237; xseA; 1.
CC HydroLase; Nuclease; Exonuclease; Complete proteome.
CC KW SEQUENCE 402 AA; 43882 MW; 145929A8372B4E08 CRC64;
CC -----
Query Match 28.3%; Score 47; DB 1; Length 402;
Best Local Similarity 47.6%; Pred. No. 18;
Matches 10; Conservative 1; Mismatches 2; Indels 8; Gaps 1;

QY 17 PAWYASRG-----IRPVG 29
||| |||
DB 89 PEWYAPRGQLSLRAAEIKPVG 109

RESULT 10
PLSC_HELPY STANDARD; PRT; 240 AA.
AC O25903;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP
DE acyltransferase) (1-AGPAT) (Lysophosphatidic acid acyltransferase)
DE (LPAAT).
GN PLSC OR HP1348.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OC NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Loftus B., Richardson J., Zhou L., Kirnes E.F., Peterson S.,
RA McKenney K., Richardson D., Dodson R., Khaklax H.G., Glodek A.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;

```

RT "The complete genome sequence of the gastric pathogen *Helicobacter pylori*."

RL Nature 388:539-547(1997).

CC -!- FUNCTION: Converts lysophosphatidic acid (LPA) into phosphatidic acid by incorporating acyl moiety at the 2 position.

CC -!- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate = CoA + 1,2-diacyl-sn-glycerol 3-phosphate.

CC -!- PATHWAY: De novo phospholipid biosynthesis; second step.

CC -!- SUBCELLULAR LOCATION: Inner membrane-associated (potential).

CC -!- SIMILARITY: Belongs to the 1-acyl-sn-glycerol-3-phosphate acyltransferase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AE000636; AAD08393.1; -.

DR PIR: D64688; D64688.

DR TIGR: HPI348; -.

DR InterPro: IPR002123; Acyltransferase.

DR InterPro: IPR004552; AGP acyltrn.

DR Pfam: PF01553; Acyltransferase; 1.

DR SMART: SM00563; PlsC; 1.

DR TIGRfams: TIGR00530; AGP acyltrn; 1.

DR Phospholipid biosynthesis; transferase; Acyltransferase;

KW Inner membrane; Complete proteome.

SQ SEQUENCE 240 AA; 27745 MW; 22BD5D0EB190BDD CRC64;

Query Match 28.0%; Score 46.5; DB 1; Length 240;

Best Local Similarity 47.6%; Pred. No. 12;

Matches 10; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 1 SRTHRSMETPTDIN-PAWY 20

Db 197 ARTRLVMELESTPTDINSPTWY 217

RESULT 11

SIR4_MOUSE STANDARD; PRT; 333 AA.

AC Q8R216; 15-MAR-2004 (Rel. 43, Created)

DT 15-MAR-2004 (Rel. 43, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE NAD-dependent deacetylase sir4uin 4 (EC 3.5.1.-) (SIR2-like protein 4).

GN SIR4 OR SIR2L4.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Colon;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klapper S., Mastrandrea A., Spies H., Wasylyuk B., Wagner L., Shenmen C.M., Schuler G.D., Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Staudt S., Schaefer S., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunatane P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fanev J., Heiton E., Kettelman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -!- FUNCTION: Probable NAD(+) + an acetylprotein = nicotinamide + O-acetyl-ADP-ribose + a protein.

CC -!- SIMILARITY: Belongs to the sir4uin family.

CC -!- SIMILARITY: Contains 1 deacetylase sir4uin-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: BC022653; AAH22653.1; ALT INIT.

DR MGD; MGI:1922637; 4930596017Rik.

DR InterPro: IPR003000; SIR2.

DR Pfam: PF02146; SIR2; 1.

DR PROSITE; PS0305; SIR2UIN; 1.

KW Hydrolase; NAD; Metal-binding; Zinc.

FT DOMAIN 42 317 DEACETYLASE SIR2UIN-TYPE.

FT ACT_SITE 158 166 BY SIMILARITY.

FT METAL 166 166 ZINC (BY SIMILARITY).

FT METAL 169 169 ZINC (BY SIMILARITY).

FT METAL 217 217 ZINC (BY SIMILARITY).

FT METAL 220 220 ZINC (BY SIMILARITY).

SQ SEQUENCE 333 AA; 37555 MW; 199FBF199617DBCO CRC64;

Query Match 28.0%; Score 46.5; DB 1; Length 333;

Best Local Similarity 34.5%; Pred. No. 18;

Matches 10; Conservative 8; Mismatches 10; Indels 1; Gaps 1;

QY 2 RTHRSMETPTDINPAWYA-SRGIRPVG 29

Db 172 QTARRVQERFOALNPWSAEACQVAPDG 200

RESULT 12

PYRG_BACAA STANDARD; PRT; 535 AA.

AC Q81JW1; 15-MAR-2004 (Rel. 43, Created)

DT 15-MAR-2004 (Rel. 43, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE CTP synthase (EC 6.3.4.2) (UTP--ammonia ligase) (CTP synthetase).

GN PYRG OR CTRA OR BA5583.

OS Bacillus anthracis (strain Ames).

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=198094;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22608414; PubMed=12721629;

RA Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T., Nelson K.E., Tetelin H., Fouts D.E., Eisen J.A., Gill S.R., Holtzapple E.K., Okstad O.A., Helgason E., Rillstone J., Wu M., Kolonay J.F., Beaman M.J., Dodson R.J., Brinkac L.M., Gwin M., DeBoy R.T., Madupu R., Daugherty S.C., Durkin H.M., Haft D.H., Nelson W.C., Peterson J.D., Pop M., Kouri H.M., Radune D., Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F., Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman J.C., Hazen A., Clime R., Redmond C., Thwaite J.E., White O., Salzberg S.L., Thomason B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B., Fraser C.M.; "The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria."

RL Nature 423:81-86(2003).

CC -!- FUNCTION: Catalyzes the ATP-dependent amination of UTP to CTP with

CC either L-glutamine or ammonia as the source of nitrogen (By
 CC similarity).
 CC -!- CATALYTIC ACTIVITY: ATP + UTP + NH(3) = ADP + phosphate + CTP.
 CC -!- ENZYME REGULATION: Allosterically activated by GTP, when glutamine
 CC is the substrate. Inhibited by CTP (By similarity).
 CC -!- PATHWAY: Pyrimidine biosynthesis; conversion of UMP to CTP; third
 CC (last) step.
 CC -!- SUBUNIT: Homotetramer (By similarity).
 CC -!- SIMILARITY: Belongs to the CTP synthase family.
 CC -!- SIMILARITY: Contains 1 type-1 glutamine amidotransferase domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE017041; AAP29225.1; -.
 CC TIGR; BAE583; -.
 CC HAVAP; MF_01227; -; 1.
 CC InterPro; IPR000991; GATase 1.
 CC InterPro; IPR004468; PyrG synth.
 CC Pfam; PF00117; GATase; 1.
 CC TIGRfams; TIGR00337; PyrG; 1.
 CC PROSITE; PS00442; GATASE_TYPE_I; 1.
 CC Pyrimidine biosynthesis; Ligase; Glutamine amidotransferase;
 CC Complete proteome.
 CC DOMAIN 1 301 AMINATOR DOMAIN.
 CC FT DOMAIN 302 535 GLUTAMINE AMIDOTRANSFERASE.
 CC FT ACT_SITE 381 381 GATASE (BY SIMILARITY).
 CC FT ACT_SITE 507 507 GATASE (BY SIMILARITY).
 CC FT ACT_SITE 509 509 GATASE (BY SIMILARITY).
 CC SEQUENCE 535 AA; 59752 MW; 0EFC670P9D77AB97 CRC64;
 CC
 CC Query Match 28.0%; Score 46.5; DB 1; Length 535;
 CC Best Local Similarity 36.7%; Pred. No. 29;
 CC Matches 11; Conservative 2; Mismatches 14; Indels 3; Gaps 1;
 CC
 CC QY 4 HRHSMETRT---PDINPAWYASRGIRPVR 30
 CC Db 459 HRHRYEFNNQFRPDMEKAGVFSGTGDGR 488
 CC
 CC RESULT 13
 CC POLN HEVBU
 CC ID POLN HEVBU STANDARD; PRT; 1693 AA.
 CC AC P29324;
 CC DT 01-DEC-1992 (Rel. 24, Created)
 CC DT 01-DEC-1992 (Rel. 24, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Non-structural polyprotein [Contains: RNA-directed RNA polymerase
 CC (EC 2.7.7.48); Helicase].
 CC OS Hepatitis E virus (strain Burma) (HEV).
 CC OC Viruses; ssRNA positive-strand viruses, no DNA stage;
 CC OC Hepatitis E-like viruses.
 CC OX NCBI_TaxID=31767;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=92024067; PubMed=1926770;
 CC RA Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
 CC RA Fry K.E., Reyes G.R.;
 CC RA "Hepatitis E virus (HEV): molecular cloning and sequencing of the
 CC full-length viral genome.";
 CC RL Virology 185:120-131(1991).
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC (RNA)(N).
 CC -!- MISCELLANEOUS: HEPATITIS E VIRUS IS THE MAJOR CAUSATIVE AGENT OF
 CC ENTERICALLY TRANSMITTED NON-A, NON-B HEPATITIS (ET-NANBH).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M73218; AAA45734.1; -.
 CC PIR; A40778; MNWVHE.
 CC MEROPS; C41.001; -.
 CC InterPro; IPR002589; Alpp.
 CC InterPro; IPR008748; Peptidase C41.
 CC InterPro; IPR001788; RNA dep RNAPol2.
 CC InterPro; IPR007095; RNA_pol_PS.
 CC InterPro; IPR007094; RNA_pol_Psvir.
 CC InterPro; IPR002588; V_methyltransf.
 CC InterPro; IPR006006; Viral_helicase1.
 CC Pfam; PF01661; Alpp; 1.
 CC Pfam; PF05417; Peptidase C41; 1.
 CC Pfam; PF00978; RNA_dep_RNAPol2; 1.
 CC Pfam; PF01443; Viral_helicase1; 1.
 CC Pfam; PF01660; Vmethyltransf; 1.
 CC Pfam; PF00506; Alpp; 1.
 CC SMART; SMO0506; Alpp; 1.
 CC Polyprotein; Transferase; RNA-directed RNA polymerase; Helicase;
 CC ATP-binding.
 CC FT NP_BIND 975 982 ATP (POTENTIAL).
 CC SQ SEQUENCE 1693 AA; 185191 MW; 2F355E463E9ED219B CRC64;
 CC
 CC Query Match 28.0%; Score 46.5; DB 1; Length 1693;
 CC Best Local Similarity 37.0%; Pred. No. 1e+02;
 CC Matches 10; Conservative 5; Mismatches 11; Indels 1; Gaps 1;
 CC
 CC QY 2 RTHRSMETRTPDINPAWY-ASRGIRP 27
 CC Db 904 RNRPGDELYLPDLAARWFANRTRP 930
 CC
 CC RESULT 14
 CC POLN HEVBU
 CC ID POLN HEVBU STANDARD; PRT; 1693 AA.
 CC AC Q04610;
 CC DT 01-OCT-1993 (Rel. 27, Created)
 CC DT 01-OCT-1993 (Rel. 27, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Non-structural polyprotein [Contains: RNA-directed RNA polymerase
 CC (EC 2.7.7.48); Helicase].
 CC OS Hepatitis E virus (strain Myanmar) (HEV).
 CC OC Viruses; ssRNA positive-strand viruses, no DNA stage;
 CC OC Hepatitis E-like viruses.
 CC OX NCBI_TaxID=31769;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=93227573; PubMed=8470371;
 CC RA Aye T.T., Uchida T., Ma M.Z., Iida F., Shikata T., Ichikawa M.,
 CC RA Rikihisa T., Winn K.;
 CC RA "Sequence and gene structure of the hepatitis E virus isolated from
 CC Myanmar.";
 CC RL Virus Genes 7:95-109(1993).
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC (RNA)(N).
 CC -!- MISCELLANEOUS: HEPATITIS E VIRUS IS THE MAJOR CAUSATIVE AGENT OF
 CC ENTERICALLY TRANSMITTED NON-A, NON-B HEPATITIS (ET-NANBH).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D10330; BAA01172.1; -.
 CC MEROPS; C41.001; -.
 CC InterPro; IPR002589; Alpp.


```
FT REPEAT 496 500 4.
FT REPEAT 504 508 5.
FT REPEAT 521 525 6.
FT REPEAT 555 559 7.
FT REPEAT 564 568 8.
FT REPEAT 573 577 9.
FT REPEAT 582 586 10.
FT REPEAT 609 613 11.
FT REPEAT 618 622 12.
FT REPEAT 627 631 13.
FT REPEAT 645 649 14.
FT REPEAT 663 667 15.
FT DOMAIN 229 326 CVS-RICH.
FT DOMAIN 374 922 PRO-RICH.
FT DOMAIN 1015 1053 ARG-RICH.
FT DOMAIN 1122 1235 CVS-RICH.
FT DOMAIN 1832 2351 PRO-RICH.
FT DOMAIN 2536 2547 GLN-RICH.
FT DOMAIN 2587 2703 PRO-RICH.
FT DOMAIN 2986 4000 GLN-RICH.
FT DOMAIN 3966 4085 PRO-RICH.
FT DOMAIN 4634 4702 PRO-RICH.
FT VARSPLIC 1 305 Missing (in isoform 2).
FT VARSPLIC 306 672 /FTid=VSP 008563.
FT HRCHEAAGGOTIRSAEHTVCSRPSPPEPGDTTDEPDA
FT LVACOGQKGGHVTSMOKERGPLOCEAKPLGKAGVLEP
FT QLEAPLNEMLPLPPEESPLSPPEESTSPPEASRLSP
FT PPEELFASLPFALHUSRLEESPLSPPEESPLSPPEESS
FT PPSPLEESPLSPPEESPLSPPEESPLSPPEESPLSPPEE
FT SPLSPPEELPSPPEESPLSPPEESPLSPPEESPLSPPEE
FT PPEASRLPPEESPLSPPEESPLSPPEESPLSPPEESPLSP
FT PMSPPPEESPLSPPEESPLSPPEESPLSPPEESPLSP
FT -> MSPPPEESPLSPPEESPLSPPEESPLSPPEESPLSP
FT PPEASRLPPEESPLSPPEESPLSPPEESPLSPPEESPLSP
FT SRLSPPEESPLSPPEESPLSPPEESPLSPPEESPLSP
FT PPDSASPPPEESPLSPPEESPLSPPEESPLSPPEESPLSP
FT HLPSPPEESPLSPPEESPLSPPEESPLSPPEESPLSP
FT PPEHLSPOAEGPHLSPOEELHLSPOEELHLSPOEELHLS
FT LSPQEEHLSPOEELHLSPOEELHLSPOEELHLSPOEELHLS
FT EKPEEPGOCAPPEELPPEESPLSPPEESPLSPPEESPLSP
FT PPLSPLPEELPSPPEESPLSPPEESPLSPPEESPLSP
FT A (in isoform 2).
FT /FTid=VSP 008559.
FT E -> EGET (in isoform 3).
FT /FTid=VSP 008560.
FT R -> H (in dbSNP:3782356).
FT /FTid=VAR 017115.
FT SQ SEQUENCE 5262 AA; 564171 MW; 2687C74CAD417E44 CRC64;
Query Match 28.0%; Score 46.5; DB 1; Length 5262;
Best Local Similarity 31.6%; Pred. No. 3.5e+02;
Matches 12; Conservative 5; Mismatches 10; Indels 11; Gaps 1;
QY 1 SRTHRSMEIRTPDINPAWYASR-----GIRP 27
Db 2037 SGTHLGLLEKLTDPVKAPLTTPRASQVPEQSPGLGRP 2074
Search completed: August 12, 2004, 14:44:52
Job time : 9.11047 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 42.7151 Seconds
(without alignments)
228.984 Million cell updates/sec

Title: US-09-700-643a-2
Perfect score: 166
Sequence: 1 SRTTHSHMEIRTPDINPAWASRGIRPVGR 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_prodot:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	152	91.6	98	Q8WN12	Q8wn12 Ovis aries
2	92	55.4	117	Q9W624	Q9w624 Carassius a
3	56.5	34.0	664	Q911W4	Q911w4 Pseudomonas
4	56	33.7	1120	Q7UPG7	Q7upg7 Rhodospirill
5	55	33.1	315	2	O87474
6	54	32.5	314	16	Q7VSA1
7	53.5	32.2	314	16	O6G348
8	53.5	32.2	314	16	Q7TWE6
9	53	31.9	327	10	Q94KU9
10	53	31.9	327	10	Q94FZ9
11	53	31.9	637	16	Q9A5E9
12	53	31.9	790	10	Q9M371
13	52	31.3	785	17	Q26276
14	51	30.7	54	4	Q7Z6Y1
15	51	30.7	450	11	Q9DA19
16	51	30.7	465	4	O60687

Query Match 91.6%; Score 152; DB 6; Length 98;
Best Local Similarity 90.0%; Pred. No. 1.4e-14;
Matches 27; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRTTHSHMEIRTPDINPAWASRGIRPVGR 30

DB 23 SRAHQSHMEIRTPDINPAWASRGIRPVGR 52

RESULT 2

Q9W624 PRELIMINARY; PRT; 117 AA.

ID Q9W624

AC Q9W624

DT 01-NOV-1999 (TREMBlrel. 12, Created)

DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

Q8w85 homo sapien
Q45883 eubacterium
P74747 synchocyst
Q7W250 bordetella
Q7W1E3 bordetella
Q89va3 bradyrhizob
Q818j6 rhodospirill
Q9gra0 hemientrot
Q8rtv7 uncultured
Q8njx4 ustilago ma
Q8fm04 corynebacte
Q8689 saccharomyc
Q8eyb3 escherichia
Q97a00 thermoplas
Q9urc6 rhodopirell
Q9f642 stigmatella
Q82ie7 streptomyce
Q8pr15 xanthomonas
Q8kz57 uncultured
Q8ywc7 anabaena sp
Q8fm14 corynebacte
Q89q2 bacteroides
Q8ukm3 agrobacteri
Q8per0 xanthomonas
Q9i022 pseudomonas
Q7wr13 bordetella
Q7w235 bordetella
Q88002 bordetella
Q45375 bordetella

ALIGNMENTS

RESULT 1

Q8WN12

ID Q8WN12 PRELIMINARY; PRT; 98 AA.

AC Q8WN12

DT 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Preprolactin-releasing peptide.

OS Ovis aries (Sheep).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Caprinae; Ovis.

OX NCBI_taxID=9940;

RN [1]

RP SEQUENCE FROM N.A.

RA Curlewis J.D., Kusters D.H.L., Barclay J.L., Anderson S.T.;

RT "Prolactin-releasing peptide (PrRP) in the ewe: cDNA cloning, mRNA

RT distribution and effects on prolactin secretion in vitro and in

RT vivo."

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF450433; AAL47178.1; -

SQ SEQUENCE 98 AA; 10513 MW; 2A53331ED62CAAB5 CRC64;

```

DE C-RF amide.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Satake H., Minakata H., Fujimoto M.;
RT "Carassius rfanide (C-RF amide)";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB020024; BAA76662.1; -.
SQ SEQUENCE 117 AA; 12879 MW; D5DC4CB22038C2B0 CRC64;

Query Match 55.4%; Score 92; DB 13; Length 117;
Best Local Similarity 52.0%; Pred. No. 11e-05;
Matches 13; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 6 HSMETPTDINPAWYASRGIRPVGR 30
Db 50 HNYDNRSPEDPPWYVGRVGRPIGR 74
:::|||||:::|||||
:::|||||:::|||||

RESULT 3
Q911W4 PRELIMINARY; PRT; 664 AA.
AC Q911W4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein PA2151.
GN PA2151.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino B., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964 (2000).
DR EMBL; AE004642; AAG05539.1; -.
DR FJX; F83376; F83376.
DR GO; GO:0003779; F-actin binding; IEA.
DR GO; GO:0004556; F-actin-amylose activity; IEA.
DR GO; GO:0005975; P-carbohydrate metabolism; IEA.
DR InterPro; IPR001589; Actbind actin.
DR InterPro; IPR006047; Alpha amyl cat.
DR Pfam; PF00128; alpha-amylose; 1.
DR PROSITE; PS00019; ACTININ 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 664 AA; 76329 MW; 8F59FEED54C308AD CRC64;

Query Match 34.0%; Score 56.5; DB 16; Length 664;
Best Local Similarity 45.8%; Pred. No. 12;
Matches 11; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 4 HRHSMETPTDINPAWYASRGIRP 27
Db 470 YRNFVFVNTPDINP-WFLQSRGP 492
:::|||||:::|||||
:::|||||:::|||||

RESULT 4
Q7UPG7 PRELIMINARY; PRT; 1120 AA.
AC Q7UPG7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN RB6951.
OS Rhodospirillum rubrum.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schleuter H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.1";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
DR EMBL; BX294145; CAD75095.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1120 AA; 123221 MW; 67C757B73421BF9B CRC64;

Query Match 33.7%; Score 56; DB 16; Length 1120;
Best Local Similarity 64.3%; Pred. No. 24;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 14 DINPAWYASRGIRP 27
Db 202 EISPAWAKYGRIP 215
:::|||||:::|||||
:::|||||:::|||||

RESULT 5
O87474 PRELIMINARY; PRT; 315 AA.
AC O87474;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Trihydroxytoluene oxygenase.
GN DNTD.
OS Burkholderia cepacia (Pseudomonas cepacia).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DNT;
RX MEDLINE=93194809; PubMed=8449889;
RA Suen W.C., Spain J.C.;
RT "Cloning and characterization of Pseudomonas sp. strain DNT genes for
RT 2,4-dinitrotoluene degradation";
RL J. Bacteriol. 175:1831-1837 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=DNT;
RX MEDLINE=99121037;
RA Haigler B.E., Johnson G.R., Suen W.C., Spain J.C.;
RT "Biochemical and genetic evidence for meta-ring cleavage of 2,4,5-
RT trihydroxytoluene in Burkholderia sp. strain DNT.";
RL J. Bacteriol. 181:965-972 (1999).
DR EMBL; AF076848; AAD12738.1; -.
DR InterPro; IPR004360; Gly_bleo_diox.
DR Pfam; PF00903; Glyoxalase; 1.
SQ SEQUENCE 315 AA; 34745 MW; E99261179022961E CRC64;

Query Match 33.1%; Score 55; DB 2; Length 315;
Best Local Similarity 44.4%; Pred. No. 8.7;
Matches 16; Conservative 3; Mismatches 11; Indels 6; Gaps 2;

QY 1 SRTHRSMETPTDINPAWYASRGIRP 30

```

"Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.";
Nature 393:537-544 (1998).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E., Kolonay J.F., Nelson W.C., Umayam L.A., Emdin M., Salzberg S.L., Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A., Bishai W.;
RA "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains.";
RT Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RL
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
CC
DR EMBL; Z95390; CAB08708.1; "-;
DR EMBL; AE007162; AAK47948.1; ALT_INIT.
DR PIR; B70569; B70569.
DR HSSP; P29132; IDFI.
DR TIGR; MT3589; -;
DR Tuberculist; RV3485C; -;
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; ADH_short.
DR Pfam; PF00106; adh_short; 1.
DR PROSITE; PS00061; ADH_SHORT; 1.
KW Hypothetical protein; Oxidoreductase; Complete proteome.
SQ SEQUENCE 314 AA; 33194 MW; 26144BA917E09274 CRC64;

Query Match 32.2%; Score 53.5; DB 16; Length 314;
Best Local Similarity 27.5%; Pred. No. 14;
Matches 11; Conservative 5; Mismatches 11; Indels 13; Gaps 1;

QY 1 SRTHR-----HSMWEIRTPDINPAWYASRGIRP 27
Db 187 SNTRWFGVGTYSVDHMMKLADELGSGSWRVNSIRP 226
||||| : : : : :
-----HSMWEIRTPDINPAWYASRGIRP 27

RESULT 8
Q7TWE6 PRELIMINARY; PRT; 314 AA.
ID Q7TWE6
AC Q7TWE6;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Probable short-chain type dehydrogenase/reductase (EC 1.-).
GN M3515C.
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1765;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=AF2122/97;
RC MEDLINE=22709107; PubMed=12788972;
RA Garner J., Eiglmeter X., Camus J.-C., Medina N., Mansoor H., Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S., Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R., Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RA "The complete genome sequence of Mycobacterium bovis.";
EL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248346; CAD95701.1; "-;
DR Oxidoreductase; Complete proteome.
SQ SEQUENCE 314 AA; 33194 MW; 26144BA917E09274 CRC64;

Query Match 32.2%; Score 53.5; DB 16; Length 314;
Best Local Similarity 27.5%; Pred. No. 14;
Matches 11; Conservative 5; Mismatches 11; Indels 13; Gaps 1;

QY 1 SRTHR-----HSMWEIRTPDINPAWYASRGIRP 27
Db 187 SNTRWFGVGTYSVDHMMKLADELGSGSWRVNSIRP 226
||||| : : : : :
-----HSMWEIRTPDINPAWYASRGIRP 27

```

RESULT 9
Q94KU9 PRELIMINARY; PRT; 327 AA.
ID Q94KU9
AC Q94KU9
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Plastid-lipid associated protein PAP1.
OS Brassica campestris (field mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3711;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21249173; PubMed=11351096;
RA Kim H.U., Wu S.S., Ratnayake C., Huang A.H.;
RT "Brassica rapa Has Three Genes That Encode Proteins Associated with
RT Different Neutral Lipids in Plastids of Specific Tissues.";
RL Plant Physiol. 126:330-341(2001).
DR EMBL: AF290563; AAK57561.1; -.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR006843; PAP_fibrillin.
DR Pfam: PF04755; PAP_fibrillin; 1.
SQ SEQUENCE 327 AA; 35678 MW; 8C7B87FCD6C02422 CRC64;

Query Match 31.9%; Score 53; DB 10; Length 327;
Best Local Similarity 43.8%; Pred. No. 18;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 4 HRHSMETPTDINPAW 19
Db 45 HRHDFKVRASDVNDWE 60

RESULT 10
Q94FZ9 PRELIMINARY; PRT; 327 AA.
ID Q94FZ9
AC Q94FZ9
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Plastid-lipid associated protein PAP1.
OS Brassica campestris (field mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3711;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21249173; PubMed=11351096;
RA Kim H.U., Wu S.S., Ratnayake C., Huang A.H.;
RT "Brassica rapa Has Three Genes That Encode Proteins Associated with
RT Different Neutral Lipids in Plastids of Specific Tissues.";
RL Plant Physiol. 126:330-341(2001).
DR EMBL: AF290566; AAK57564.1; -.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR006843; PAP_fibrillin.
DR Pfam: PF04755; PAP_fibrillin; 1.
SQ SEQUENCE 327 AA; 35644 MW; 6116E7F1B6C02C88 CRC64;

Query Match 31.9%; Score 53; DB 10; Length 327;
Best Local Similarity 43.8%; Pred. No. 18;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 4 HRHSMETPTDINPAW 19
Db 45 HRHDFKVRASDVNDWE 60

RESULT 11

```

```

Q9A5E9 PRELIMINARY; PRT; 637 AA.
ID Q9A5E9
AC Q9A5E9
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Sensory box histidine kinase/response regulator.
GN CC2501.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Sait J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -!- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
CC -!- SIMILARITY: TO PROKARYOTE SENSORY TRANSDUCTION PROTEINS.
DR EMBL: AE005918; AAK24472.1; -.
DR PIR: D87559; D87559.
DR HSP: C56312; 3TMV.
DR TIGR: CC2501; -.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0016301; F:kinase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0000156; F:two-component response regulator activity; IEA.
DR GO: GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO: GO:0007600; P:sensory perception; IEA.
DR GO: GO:0000160; P:two-component signal transduction system (p. . .); IEA.
DR InterPro: IPR003594; ATPbind_Atpase.
DR InterPro: IPR004358; Bact_sens_pr_C.
DR InterPro: IPR005467; His_kinase.
DR InterPro: IPR003661; His_kin_N.
DR InterPro: IPR001610; PAC.
DR InterPro: IPR000700; PAS-assoc_C.
DR InterPro: IPR000014; PAS_domain.
DR InterPro: IPR001789; Response_reg.
DR Pfam: PF02518; HATPase_c; 1.
DR Pfam: PF00512; HSKA; 1.
DR Pfam: PF00785; PAC; 1.
DR Pfam: PF00989; PAS; 1.
DR Pfam: PF00072; response_reg; 1.
DR PRINTS: PR00344; BCTRSENSOR.
DR ProDom: PD000039; Response_reg; 1.
DR SMART: SM00387; HATPase_c; 1.
DR SMART: SM00388; HSKA; 1.
DR SMART: SM00086; PAC; 1.
DR SMART: SM00091; PAS; 1.
DR SMART: SM00448; REC; 1.
DR TIGRfams: TIGR00229; sensory_box; 1.
DR PROSITE: PS50109; HIS_KIN; 1.
DR PROSITE: PS50113; PAC; 1.
DR PROSITE: PS50110; RESPONSE_REGULATORY; 1.
DR PROSITE: PS50110; RESPONSE_REGULATORY; 1.
KW Kinase; Phosphorylation; Sensory transduction; Transferase;
SQ Complete proteome.
SQ SEQUENCE 637 AA; 68511 MW; 0EDEAF76FFA8611 CRC64;

Query Match 31.9%; Score 53; DB 16; Length 637;
Best Local Similarity 48.0%; Pred. No. 36;
Matches 12; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

```

```

QY 4 HRHSMETPTDINPAWASRGIRPV 28
DB 22 HRDSDLRSPAINPAIRVILRAV 46

RESULT 12
Q9N371 PRELIMINARY; PRT; 790 AA.
AC Q9N371;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN F15G16.60.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RN SEQUENCE FROM N.A.
RP De Haan M., Maarse A.C., Griwell L.A., Mewes H.W., Lemcke K.,
RA Mayer K.F.X., Quettier F., Salanoubat M.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP EU Arabidopsis sequencing project;
RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AL132959; CAB71097.1; -.
DR PIR; T47959; T47959.
KW Hypothetical protein.
SQ SEQUENCE 790 AA; 87376 MW; B222724B75690F30 CRC64;

Query Match 31.9%; Score 53; DB 10; Length 790;
Best Local Similarity 55.6%; Pred. No. 46;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 13 PDINPAWASRGIRPVGR 30
DB 366 PPHNPTVGSRLQPHGR 383

RESULT 13
O26276 PRELIMINARY; PRT; 785 AA.
AC O26276;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sensory transduction histidine kinase.
GN MTH174.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwni N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT delcAH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155 (1997).
DR EMBL; AE008085; AB84680.1; -.
DR PIR; F69099; F69099.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.

```

```

DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.
DR InterPro; IPR003594; ATPbind_Alpase.
DR InterPro; IPR000515; BPD_transp.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR001610; PAC_
DR InterPro; IPR000700; PAS-assoc C.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF02518; HAIPase_c; 1.
DR Pfam; PF00785; PAC; 1.
DR Pfam; PF00989; PAS; 3.
DR SMART; SM00387; HAIPase_c; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAC; 3.
DR TIGRfams; TIGR00229; sensory box; 3.
DR PROSITE; PS00402; BPD_TRANSF_INN_MEMBER; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
DR PROSITE; PS50113; PAC; 1.
DR PROSITE; PS50112; PAS; 3.
KW Complete proteome.
SQ SEQUENCE 785 AA; 87726 MW; ADD502C928307986 CRC64;

Query Match 31.3%; Score 52; DB 17; Length 785;
Best Local Similarity 46.4%; Pred. No. 64;
Matches 13; Conservative 1; Mismatches 12; Indels 2; Gaps 1;

QY 4 HRHSMETPTDINPAW--YASRGIRPVG 29
DB 412 HRATPRRRPGNYRWVEYVDRPIRTDG 439

RESULT 14
Q7Z6V1 PRELIMINARY; PRT; 54 AA.
AC Q7Z6V1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DJ479J7.3 (Sushi-repeat protein (SRPUL)) (Fragment).
GN DJ479J7.3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RP Lawlor S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL035608; CAB55682.1; -.
FT NON_TER 54
SQ SEQUENCE 54 AA; 6110 MW; E2F3C39P7B961A9F CRC64;

Query Match 30.7%; Score 51; DB 4; Length 54;
Best Local Similarity 50.0%; Pred. No. 5;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 12 TPDINPAWASRGIRP 27
DB 18 TPAVTPTWAGSGYYP 33

RESULT 15
Q9DA19 PRELIMINARY; PRT; 450 AA.
AC Q9DA19;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 1700023B02Rik protein.
GN CIR OR 1700023B02RIK.

```

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.,
 RA "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL; AK006260; BAB24488.1; -.
 DR MGD; MGI:1914185; Ctr.
 DR GO; GO:0005489; P:electron transporter activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR000345; CytC_heme_BS.
 DR PROSITE; PS00190; CYTOCHROME_C; 1.
 SQ SEQUENCE 450 AA; 51852 MW; F32F11BE6D6A4EAC CRC64;

Query Match 30.7%; Score 51; DB 11; Length 450;
 Best Local Similarity 46.7%; Pred.No 49;
 Matches 14; Conservative 2; Mismatches 8; Indels 6; Gaps 1;
 Qy 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30
 ||:|||||:|
 Db 373 SRSHRSPEKKGSDRN-----RGIRSR 396

Search completed: August 12, 2004, 14:49:01
 Job time : 43.7151 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: August 12, 2004, 14:37:35 ; Search time 65.7849 Seconds
(without alignments)
133.146 Million cell updates/sec

Title: US-09-700-643A-2
Perfect score: 166
Sequence: 1 SRTTHSHMEIRTPDINPAVASEGIRPVGRX 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1980s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	165	99.4	30	3 AAY49299	Aay49299 19p2 liga
2	165	99.4	31	2 AAW31371	Aaw31371 Human typ
3	165	99.4	31	2 AAW87615	Aaw87615 Human 19p
4	165	99.4	31	2 AAW97235	Aaw97235 Human typ
5	165	99.4	31	3 AAY49291	Aay49291 19p2 liga
6	165	99.4	31	3 AAB10362	Aab10362 Human oxy
7	165	99.4	31	4 AAB90991	Aab90991 Prolactin
8	165	99.4	31	4 AAG62531	Aag62531 Human CRH
9	165	99.4	31	5 AAE26401	Aae26401 Human prr
10	165	99.4	31	6 ABU60843	Abu60843 Peptide p
11	165	99.4	31	6 ABU60827	Abu60827 Peptide p
12	165	99.4	31	7 ADC71228	Adc71228 Human pep
13	165	99.4	32	2 AAW31392	Aaw31392 Human typ
14	165	99.4	32	3 AAB10363	Aab10363 Human oxy
15	165	99.4	32	4 AAG62532	Aag62532 Human CRH
16	165	99.4	32	6 ABU60844	Abu60844 Peptide p
17	165	99.4	33	2 AAW31393	Aaw31393 Human typ
18	165	99.4	33	3 AAB10364	Aab10364 Human oxy
19	165	99.4	33	4 AAG62533	Aag62533 Human CRH
20	165	99.4	33	6 ABU60845	Abu60845 Peptide p
21	165	99.4	37	2 AAW31390	Aaw31390 Human typ
22	165	99.4	87	2 AAW97226	Aaw97226 Human typ
23	165	99.4	87	3 AAB10361	Aab10361 Human oxy
24	165	99.4	87	4 AAG62530	Aag62530 Human CRH
25	156	94.0	31	4 AAB90995	Aab90995 Prolactin

26	152	91.6	31	2 AAW31371	Aaw31371 Bovine G
27	152	91.6	31	2 AAW95188	Aaw95188 Bovine pi
28	152	91.6	31	2 AAW87613	Aaw87613 Bovine 19
29	152	91.6	31	2 AAW97218	Aaw97218 Bovine pi
30	152	91.6	31	3 AAY49298	Aay49298 19p2 liga
31	152	91.6	31	3 AAY49290	Aay49290 19p2 liga
32	152	91.6	31	3 AAB10347	Aab10347 Bovine ox
33	152	91.6	31	4 AAG62516	Aag62516 Bovine CR
34	152	91.6	31	5 AAE26399	Aae26399 Bovine pr
35	152	91.6	31	6 ABU60825	Abu60825 Peptide p
36	152	91.6	31	6 ABU60831	Abu60831 Peptide p
37	152	91.6	32	2 AAW31372	Aaw31372 Bovine G
38	152	91.6	32	2 AAW95189	Aaw95189 Bovine pi
39	152	91.6	32	3 AAB10348	Aab10348 Bovine ox
40	152	91.6	32	4 AAG62517	Aag62517 Bovine CR
41	152	91.6	32	6 ABU60832	Abu60832 Peptide p
42	152	91.6	33	2 AAW31373	Aaw31373 Bovine G
43	152	91.6	33	2 AAW95190	Aaw95190 Bovine pi
44	152	91.6	33	3 AAY49297	Aay49297 19p2 liga
45	152	91.6	33	3 AAB10349	Aab10349 Bovine ox

ALIGNMENTS

RESULT 1
AAY49299
ID AAY49299 standard; peptide; 30 AA.
XX
AC AAY49299;
XX
DT 22-FEB-2000 (first entry)
XX
DE 19P2 ligand peptide fragment.
XX
KW Monoclonal antibody; 19p2 ligand; diagnosis; prolactin secretion;
KW pituitary; regulatory mechanism; central nervous system; pancreatic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 30
FT /note= "C-terminal amide"
XX
PN WO9960112-A1.
XX
PD 25-NOV-1999.
XX
PF 20-MAY-1999; 99WO-JP002650.
XX
PR 21-MAY-1998; 98JP-00140293.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Matsumoto H, Kitada C, Hinuma S;
XX
PS WPI; 2000-039381/03.
XX
DR
XX
PT New monoclonal antibodies, useful in diagnosis, as drugs and in studying diseases related to ligand abnormality.
XX
PS Disclosure; Page 27; 73pp; Japanese.
XX
CC The invention provides a monoclonal antibody which has a specific reaction with the part peptide of the C-terminal of 19p2 ligand or its derivative. The antibodies can be used in diagnosis or to treat or prevent diseases associated with abnormality in the pituitary function regulatory mechanism (e.g. promotion of prolactin secretion), central nervous regulatory mechanism, and pancreatic function regulatory mechanism. The antibody-based immunoassay can also be applied in clarifying the physiological functions of the ligand and its derivative.
XX
CC Sequences AAY49290-302 represent peptide fragments of the 19p2 ligand

SQ Sequence 30 AA;		Best Local Similarity 100.0%; Pred. No. 5.4e-18;		Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Query Match		99.4%; Score 165; DB 3; Length 30;			
Best Local Similarity 100.0%; Pred. No. 5.2e-18;					
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1 SRTHSHMEIRTPDINPAWYASRGIRPVGR 30				
DB	1 SRTHSHMEIRTPDINPAWYASRGIRPVGR 30				
RESULT 2					
AAW31391	ID AAW31391 standard; peptide; 31 AA.				
XX					
AC	AAW31391;				
XX					
DT	06-APR-1998 (first entry)				
XX					
DE	Human type G protein-coupled receptor ligand fragment 1.				
XX					
XX	G protein-coupled receptor; ligand binding; pharmaceutical; modulator;				
KW	pituitary; central nervous system; pancreas; prophylactic;				
KW	therapeutic agent.				
XX					
OS	Homo sapiens.				
XX					
PN	W09724436-A2.				
XX					
PD	10-JUL-1997.				
XX					
PF	26-DEC-1996; 96WO-JP003821.				
XX					
PR	28-DEC-1995; 95JP-00343371.				
PR	15-MAR-1996; 96JP-00059419.				
PR	12-AUG-1996; 96JP-00211805.				
PR	18-SEP-1996; 96JP-00245573.				
XX					
PA	(TAKE) TAKEDA CHEM IND LTD.				
XX					
PI	Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;				
PI	Kitada C;				
XX					
DR	WPI: 1997-363672/33.				
DR	N-PSDB; AAV02428.				
XX					
PT	Ligand peptide for G protein-coupled receptor - acts by modulating				
PT	function in the central nervous system, pancreas and pituitary gland.				
XX					
PS	Claim 2; Page 184; 258pp; English.				
XX					
CC	This sequence represents a peptide fragment from a novel human type				
CC	ligand polypeptide corresponding to amino acid residues 23 to 53 of the				
CC	sequence represented in AAW31390 and is used in an assay to monitor				
CC	ligand binding to the G protein-coupled receptor protein. Pharmaceutical				
CC	compositions containing this ligand may be used as a pituitary function				
CC	modulator, a central nervous system modulator or a pancreatic function				
CC	modulator. This ligand could have specific applications as a prophylactic				
CC	or therapeutic agent for dementia, depression, hyperkinetic syndrome,				
CC	disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,				
CC	growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,				
CC	hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,				
CC	cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,				
CC	rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,				
CC	amyotrophic lateral sclerosis, acute myocardial infarction, infertility,				
CC	spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,				
CC	osteoporosis and/or oligogalactia. Assays can also be developed to screen				
CC	compounds which are capable of altering the binding activity of the				
CC	ligand affecting activation of the G protein-coupled receptor protein				
XX					
SQ	Sequence 31 AA;				
Query Match		99.4%; Score 165; DB 2; Length 31;			
Best Local Similarity 100.0%; Pred. No. 5.4e-18;					
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1 SRTHSHMEIRTPDINPAWYASRGIRPVGR 30				
DB	1 SRTHSHMEIRTPDINPAWYASRGIRPVGR 30				
RESULT 3					
AAW87615	ID AAW87615 standard; peptide; 31 AA.				
XX					
AC	AAW87615;				
XX					
DT	29-MAR-1999 (first entry)				
XX					
DE	Human 19P2 ligand.				
XX					
KW	19P2 ligand; G protein coupled receptor; pituitary;				
KW	prolactin releasing peptide; human; dementia; breast cancer; therapy.				
OS	Homo sapiens.				
XX					
PN	EP897417-A2.				
XX					
PD	30-DEC-1998.				
XX					
PF	25-JUN-1998; 98EP-00111725.				
XX					
PR	27-JUN-1997; 97JP-00172118.				
XX					
PA	(TAKE) TAKEDA CHEM IND LTD.				
XX					
PI	Suenaga M, Moriya T, Tanaka Y, Nishimura O;				
XX					
DR	WPI: 1999-047884/05.				
XX					
PT	Producing a 19P2 pituitary G protein receptor ligand - by cleavage of a				
PT	fusion protein, useful for preventing and treating dementia, breast				
PT	cancer, renal failure and autoimmune disease.				
XX					
PS	Claim 5; Page 35; 56pp; English.				
XX					
CC	This is the amino acid sequence of the human pituitary G protein-coupled				
CC	receptor ligand 19P2L. A method suitable for commercial high-level				
CC	production of 19P2L comprises expressing the ligand in host cells as a				
CC	recombinant fusion protein e.g. with human basic fibroblast growth factor				
CC	(see AAV83796-97) that has been modified to include an N-terminal				
CC	cysteine residue. The ligand is released from the fusion by cyanylation				
CC	followed by ammonolysis. 19P2L has prolactin secretion-stimulating and				
CC	(at high doses) prolactin secretion-inhibiting properties. It can be used				
CC	in the treatment and prevention of various diseases including: senile				
CC	dementia, cerebrovascular dementia, and dementia associated with:				
CC	neurological disorders (e.g. Alzheimer's disease, Parkinson's disease,				
CC	Pick's disease, Huntington's disease), infectious diseases (e.g.				
CC	Creutzfeldt-Jakob's), endocrine or metabolic disease or toxicosis (e.g.				
CC	hypothyroidism, vitamin B12 deficiency, alcoholism, intoxication by				
CC	drugs, metal and organic compounds), tumorigenic diseases (e.g. brain				
CC	tumour), traumatic diseases (e.g. chronic subarachnoid haemorrhage, and				
CC	other types of dementia, depression, hyperactive child syndrome				
CC	(microencephalopathy) and disturbance of consciousness. It is also useful				
CC	for prevention and treatment of diseases associated with prolactin hypo				
CC	and hypersecretion respectively, including: hyperprolactinaemia				
CC	disease (hypersecretion disorders), and seminal vesicle hypoplasia,				
CC	osteoporosis, menopausal syndrome and renal failure (hyposecretion				
CC	disorders). The 19P2 polypeptide/amide is also useful as a test reagent				
CC	for study of the prolactin secretory function or as a lactagogue in				
CC	mammalian farm animals				
XX					
SQ	Sequence 31 AA;				
Query Match		99.4%; Score 165; DB 2; Length 31;			

Best Local Similarity 100.0%; Pred. No. 5.4e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTTHSMSEIRTPDINPAWYASRGIRPVGR 30
DB 1 SRTTHSMSEIRTPDINPAWYASRGIRPVGR 30

RESULT 4
AAW97235
ID AAW97235 standard; peptide; 31 AA.
XX
AC AAW97235;
XX
DT 06-MAY-1999 (first entry)
XX
DE Human type ligand polypeptide fragment.
XX
KW Rat type ligand; modulation; prolactin secretion;
KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacosgenesis;
KW menopausal syndrome; euthyroid; hypometabolism; lactation;
KW pituitary adenomatosis; brain tumour; emeniorrhea; autoimmune disease;
KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;
KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyscoospermia;
KW contraception; placental function; choriocarcinoma; hydatid mole;
KW irruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;
KW abnormal lipidmetabolism; oxytocia.
XX
OS Homo sapiens.
XX
FN WO9858962-A1.
XX
PD 30-DEC-1998.
XX
XX 22-JUN-1998; 98WO-JP002765.
XX
PR 23-JUN-1997; 97JP-00165437.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Hinuma S, Kawamata Y, Fujii R, Matsumoto H;
XX WPI; 1999-105614/09.
XX
XX Use of G protein-coupled receptor ligands - for modulating prolactin
PT secretion or placental function, e.g. for treating menopausal syndrome,
PT tumours, autoimmune disease or abnormal pregnancy.
XX
PS Claim 3; Page 159; 241pp; English.
XX
CC The present sequence represents a human type ligand fragment. It is used
CC in the course of the invention. The specification describes an agent for
CC modulating prolactin secretion which comprises a ligand polypeptide or a
CC salt, for a G protein-coupled receptor (GPCR) protein. The agents for
CC promoting prolactin secretion can be used for treating or preventing
CC hypovarianism, gonocyst cacosgenesis, menopausal syndrome, euthyroid or
CC hypometabolism. They can be used for promoting lactation in a domestic
CC mammal and as an aphrodisiac. The agents for inhibiting prolactin
CC secretion can be used for treating or preventing pituitary adenomatosis,
CC brain tumour, emeniorrhea, autoimmune disease, prolactinoma,
CC infertility, impotence, amenorrhea, galactorrhea, acromegaly, Chiari-
CC Frommel syndrome, Argonz-del Castillo syndrome, Forbes-Albright syndrome,
CC lymphoma, Sheehan syndrome or dyscoospermia. The inhibitory agents can
CC also be used as contraceptives. The agents for modulating placental
CC function can be used for treating or preventing choriocarcinoma, hydatid
CC mole, irruption mole, abortion, unthrifty fetus, abnormal
CC saccharometabolism, abnormal lipidmetabolism or oxytocia
XX
SQ Sequence 31 AA;

Query Match 99.4%; Score 165; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 5.4e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTTHSMSEIRTPDINPAWYASRGIRPVGR 30
DB 1 SRTTHSMSEIRTPDINPAWYASRGIRPVGR 30

RESULT 5
AAAY49291
ID AAY49291 standard; peptide; 31 AA.
XX
AC AAY49291;
XX
DT 22-FEB-2000 (first entry)
XX
DE 19P2 ligand peptide fragment.
XX
KW Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;
KW pituitary; regulatory mechanism; central nervous system; pancreatic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 31
FT /note= "C-terminal amide"
XX
XX WO9960112-A1.
XX
XX 25-NOV-1999.
XX
XX 20-MAY-1999; 99WO-JP002650.
XX
PR 21-MAY-1998; 98JP-00140293.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Matsumoto H, Kitada C, Hinuma S;
XX WPI; 2000-039381/03.
XX
XX New monoclonal antibodies, useful in diagnosis, as drugs and in studying
PT diseases related to ligand abnormality.
XX
PS Disclosure; Page 26; 73pp; Japanese.
XX
CC The invention provides a monoclonal antibody which has a specific
CC reaction with the part peptide of the C-terminal of 19P2 ligand or its
CC derivative. The antibodies can be used in diagnosis or to treat or
CC prevent diseases associated with abnormality in the pituitary function
CC regulatory mechanism (e.g. promotion of prolactin secretion), central
CC nervous regulatory mechanism, and pancreatic function regulatory
CC mechanism. The antibody-based immunoassay can also be applied in
CC clarifying the physiological functions of the ligand and its derivative.
CC Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand
XX
SQ Sequence 31 AA;

Query Match 99.4%; Score 165; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 5.4e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTTHSMSEIRTPDINPAWYASRGIRPVGR 30
DB 1 SRTTHSMSEIRTPDINPAWYASRGIRPVGR 30

RESULT 6
AAB10362
ID AAB10362 standard; peptide; 31 AA.
XX
AC AAB10362;
XX
DT 24-NOV-2000 (first entry)

XX DE Human oxytocin secretion promoting peptide SEQ ID NO: 32.
 XX DE Human; oxytocin secretion promoter; G protein-coupled receptor protein;
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;
 KW veterinary medicine; milk production.
 XX OS Homo sapiens.
 XX FN WO200038704-A1.
 XX PD 06-JUL-2000.
 XX PP 22-DEC-1999; 99WO-JP007199.
 XX PR 25-DEC-1998; 98JP-00369565.
 XX PA (TAKE) TAKEDA CHEM IND LTD.
 XX PI Matsumoto H, Kitada C, Hinuma S;
 XX WPI; 2000-452298/39.
 XX PT Physiologically-active polypeptide recognized as ligand by G protein-
 PT coupled receptor protein, for promoting secretion of oxytocin, as drugs
 PT for diseases relating to oxytocin secretion and in veterinary medicine.
 XX PS Disclosure; Page 62; 72pp; Japanese.
 XX CC This invention describes a novel oxytocin secretion-regulating agent
 CC which contains a ligand peptide or its salt for the G protein-coupled
 CC receptor protein. It is useful in the form of drugs for ameliorating,
 CC preventing and treating diseases relating to oxytocin secretion e.g. weak
 CC pains and atonic bleeding, before and after expulsion of placenta,
 CC uterine recovery failure, caesarean section, stoppage of artificial
 CC fertilization or galactostasis and is also applicable in veterinary
 CC medicine for promoting milk production in cow, goat and pig. This
 CC sequence represents a human peptide which acts as an oxytocin secretion
 CC promoter
 XX SQ Sequence 31 AA;
 Query Match 99.4%; Score 165; DB 3; Length 31;
 Best Local Similarity 100.0%; Pred. No. 5.4e-18;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30
 DB 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30
 RESULT 7
 AAB90991
 ID AAB90991 standard; peptide; 31 AA.
 AC AAB90991;
 XX 22-JUN-2001 (first entry)
 XX Prolactin releasing peptide SEQ ID NO:165.
 DE DE Protection; endogenous therapeutic peptide; peptidase; conjugation;
 KW blood component; modification; succinimide; maleimide group; amino;
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
 XX OS Homo sapiens.
 XX OS Synthetic.
 XX PN WO200069900-A2.
 XX PD 23-NOV-2000.
 XX (TAKE) TAKEDA CHEM IND LTD.

PF 17-MAY-2000; 2000WO-US013576.
 XX 17-MAY-1999; 99US-0134406P.
 PR 10-SEP-1999; 99US-0153406P.
 PR 15-OCT-1999; 99US-0159783P.
 XX (CONJ-) CONJUCHEM INC.
 XX PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
 XX WPI; 2001-112059/12.
 XX PT Modifying and attaching therapeutic peptides to albumin prevents
 PT peptidase degradation, useful for increasing length of in vivo activity.
 XX PS Disclosure; Page 244; 733pp; English.
 XX CC The present invention describes a modified therapeutic peptide (I)
 CC comprising a therapeutically active amino acid region (III) and a
 CC reactive group (II) (e.g. succinimide and maleimide groups) attached to
 CC a less therapeutically active amino acid region (IV), which covalently
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
 CC factors and neurotransmitters, to protect them from peptidase activity in
 CC vivo for the treatment of various disorders. Endogenous therapeutic
 CC peptides are not suitable as drug candidates as they require frequent
 CC administration due to rapid degradation by peptidases in the body.
 CC Modifying and attaching therapeutic peptides to albumin prevents or
 CC reduces the action of peptidases to increase length of activity (half
 CC life) and specificity as bonding to large molecules decreases
 CC intracellular uptake and interference with physiological processes.
 CC AAB90929 to AAB92441 represent peptides which can be used in the
 CC exemplification of the present invention
 XX SQ Sequence 31 AA;
 Query Match 99.4%; Score 165; DB 4; Length 31;
 Best Local Similarity 100.0%; Pred. No. 5.4e-18;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30
 DB 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30
 RESULT 8
 AAG62531
 ID AAG62531 standard; peptide; 31 AA.
 AC AAG62531;
 XX 24-AUG-2001 (first entry)
 XX Human CRH releasing protein related peptide SEQ ID NO: 32.
 DE DE Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;
 KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;
 KW Addison's disease; adrenal gland hyperfunction; obesity.
 XX OS Homo sapiens.
 XX PN WO200135984-A1.
 XX PD 25-MAY-2001.
 XX PF 17-NOV-2000; 2000WO-JP008119.
 XX PR 18-NOV-1999; 99JP-00327900.
 PR 26-SEP-2000; 2000JP-00297073.
 XX (TAKE) TAKEDA CHEM IND LTD.

PI Kitada C, Matsumoto H, Hinuma S;
 XX WPI; 2001-355552/37.
 XX
 XX Use of G protein receptor ligand or peptide for controlling corticotropin
 PT releasing hormone secretion.
 XX
 XX Claim 3; Page 73-74; 90pp; Japanese.
 PS
 XX The present sequence describes a method of controlling the secretion of
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein
 CC receptor ligand. This can be used to control the secretion of CRH and is
 CC useful as an analgesic or for treating, preventing or ameliorating
 CC diseases associated with CRH secretion such as hyperaldosteronism,
 CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The
 CC present sequence is a peptide used in the exemplification of the
 CC invention
 XX
 XX Sequence 31 AA;
 SQ
 Query Match 99.4%; Score 165; DB 4; Length 31;
 Best Local Similarity 100.0%; Pred. No. 5.4e-18;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRTHSHMEIRTPDINPAWYASRGIRPVGR 30
 DB 1 SRTHSHMEIRTPDINPAWYASRGIRPVGR 30
 RESULT 9
 AAE26401
 ID AAE26401 standard; peptide; 31 AA.
 XX
 AC AAE26401;
 XX
 XX 13-DEC-2002 (first entry)
 DT
 XX Human PrRP-31 peptide.
 DE
 XX Human; wakefulness; sleep disorder; prolactin releasing peptide receptor;
 KW PrRP; GPR10; therapy; epilepsy; narcolepsy; sleepiness; sleep apnoea;
 KW insomnia; idiopathic hypersomnia; psychogenic hypersomnia; seizure;
 KW anticonvulsant.
 XX
 XX Homo sapiens.
 OS
 XX US2002037533-A1.
 PN
 XX 28-MAR-2002.
 PD
 XX 17-AUG-2001; 2001US-00932161.
 PF
 XX 28-APR-2000; 2000US-00560915.
 PR
 XX (CIVE/) CIVELLI O.
 PA (LINS/) LIN S.
 PI Civelli O, Lin S;
 XX
 XX WPI; 2002-403931/43.
 DR
 XX Screening for compounds useful for promoting wakefulness or sleep, and
 PT for treating sleeping disorders, e.g. insomnia, hypersomnia or sleep
 PT apnea, comprises administering a prolactin releasing peptide agonist or
 PT antagonist.
 XX
 XX Disclosure; Page 24; 35pp; English.
 PS
 XX The present invention relates to a method of screening for compounds for
 CC promoting wakefulness or sleep in a mammal. The method involves
 CC administering a prolactin releasing peptide (PrRP) receptor (GPR10)

CC agonist or antagonist respectively and determining the ability of the
 CC compound to promote wakefulness or sleep. The compounds identified from
 CC the method are used in the therapy of epilepsy and other diseases
 CC associated with absence seizures and in promoting wakefulness and sleep
 CC in individuals having sleep disorders such as insomnia and narcolepsy.
 CC PrRP receptor agonists may be used to treat common disorders which lead
 CC to sleepiness, e.g. sleep apnoea, narcolepsy, idiopathic hypersomnia and
 CC psychogenic hypersomnia. PrRP receptor antagonists are useful for
 CC promoting sleep and for treating insomnia such as adjustment sleep
 CC disorder and psychophysiological insomnia. The present sequence is human
 CC PrRP-31 peptide
 XX
 XX Sequence 31 AA;
 SQ
 Query Match 99.4%; Score 165; DB 5; Length 31;
 Best Local Similarity 100.0%; Pred. No. 5.4e-18;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRTHSHMEIRTPDINPAWYASRGIRPVGR 30
 DB 1 SRTHSHMEIRTPDINPAWYASRGIRPVGR 30
 RESULT 10
 ABU60843
 ID ABU60843 standard; peptide; 31 AA.
 XX
 AC ABU60843;
 XX
 XX 06-MAY-2003 (first entry)
 DT
 XX Peptide production by gene recombination associated peptide #27.
 DE
 XX Peptide production; low-molecular peptide; KiSS-1; GPR8 ligand;
 KW Gene recombination.
 KW
 XX Homo sapiens.
 OS
 XX WO200292829-A1.
 PN
 XX 21-NOV-2002.
 PD
 XX 16-MAY-2002; 2002WO-JP004735.
 PF
 XX 17-MAY-2001; 2001JP-00147341.
 PR
 XX (TAKE) TAKEDA CHEM IND LTD.
 PA
 XX Nishimura O, Suenaga M, Ito T, Kitada C;
 PI
 XX WPI; 2003-129302/12.
 DR
 XX Process for producing peptides e.g. KiSS-1 peptide and GPR8 ligand for
 PT subsequent applications by gene recombination technique through tandem
 PT repeats to provide precursor protein with specific cleavage sites.
 XX
 XX Disclosure; Page 67; 87pp; Japanese.
 PS
 XX The invention describes a method of producing a peptide comprising the
 CC excision of the N and C-terminals of a target peptide with enzymes or
 CC chemically through the attached cleavage sites repeated by ligation in a
 CC precursor protein. The method is for producing (low-molecular) peptides
 CC e.g. KiSS-1 peptide and GPR8 ligand for subsequent applications by the
 CC gene recombination technique through tandem repeats to provide a
 CC precursor protein with specific cleavage sites. With this method, peptide
 CC production can be carried out easily to provide large quantities of the
 CC required peptides. This is the amino acid sequence of a peptide
 CC associated with the peptide production method of the invention
 XX
 XX Sequence 31 AA;
 SQ
 Query Match 99.4%; Score 165; DB 6; Length 31;
 Best Local Similarity 100.0%; Pred. No. 5.4e-18;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRTHSHMEIRTPDINPAWYASRGIRPVGR 30
Db 1 SRTHSHMEIRTPDINPAWYASRGIRPVGR 30

RESULT 11
ABU60827
ID ABU60827 standard; peptide; 31 AA.
AC ABU60827;
XX
DT 06-MAY-2003 (first entry)
XX
XX Peptide production by gene recombination associated peptide #11.
XX
XX Peptide production; low-molecular peptide; KISS-1; GPR8 ligand;
KW gene recombination.
KW
XX Homo sapiens.
OS
XX WO200292829-A1.
XX
XX 21-NOV-2002.
XX
XX 16-MAY-2002; 2002WO-JP004735.
XX
XX 17-MAY-2001; 2001JP-00147341.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Nishimura O, Suenaga M, Ito T, Kitada C;
XX
XX WPI; 2003-129302/12.
XX
XX Process for producing peptides e.g. Kiss-1 peptide and GPR8 ligand for
PT subsequent applications by gene recombination technique through tandem
PT repeats to provide precursor protein with specific cleavage sites.
XX
XX Disclosure; Page 59; 87pp; Japanese.
XX
XX The invention describes a method of producing a peptide comprising the
CC excision of the N and C-terminals of a target peptide with enzymes or
CC chemically through the attached cleavage sites repeated by ligation in a
CC precursor protein. The method is for producing (low-molecular) peptides
CC e.g. Kiss-1 peptide and GPR8 ligand for subsequent applications by the
CC gene recombination technique through tandem repeats to provide a
CC precursor protein with specific cleavage sites. With this method, peptide
CC production can be carried out easily to provide large quantities of the
CC required peptides. This is the amino acid sequence of a peptide
CC associated with the peptide production method of the invention
XX
SQ Sequence 31 AA;

Query Match 99.4%; Score 165; DB 6; Length 31;
Best Local Similarity 100.0%; Pred. No. 5.4e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRTHSHMEIRTPDINPAWYASRGIRPVGR 30
Db 1 SRTHSHMEIRTPDINPAWYASRGIRPVGR 30

RESULT 12
ADCT71228
ID ADCT71228 standard; peptide; 31 AA.
XX
XX ADCT71228;
XX
XX 18-DEC-2003 (first entry)
DT
XX
XX Human peptide sequence 2 related to the human serine protease.

XX
KW human; serine protease precursor; hormone; neurohypophyseal; diabetes;
KW diabetic retinopathy; cataract; antidiabetic; ophthalmological.
XX
OS Homo sapiens.
XX
XX WO2003062429-A1.
XX
XX 31-JUL-2003.
XX
XX 22-JAN-2003; 2003WO-JP000547.
XX
XX 23-JAN-2002; 2002JP-00013849.
PR
XX 10-OCT-2002; 2002JP-00298003.
XX
XX (YAMA) YAMANOUCHI PHARM CO LTD.
XX
XX Kagoshima M, Yamaji N, Takeda M, Abe K, Kawabe T;
PI
XX WPI; 2003-598754/56.
XX
XX Protease precursor for developing treatment for diabetes and cataracts.
PT
XX Example 10; SEQ ID NO 13; 61pp; Japanese.
XX
XX This invention relates to a novel human serine protease precursor
CC polypeptide. Specifically, it refers to a novel type II transmembrane
CC serine protease that participates in the control of hormones produced by
CC the pancreatic gland. As such, it can be used in a screening method for
CC the identification of compounds that are useful for the treatment of
CC diseases where a neurohypophyseal hormone participates, such as
CC diabetes, diabetic retinopathy and cataract. The present invention
CC describes the compounds of this invention as antidiabetic and
CC ophthalmological. This peptide sequence is human peptide 2 related to the
CC human serine protease of the invention.
XX
SQ Sequence 31 AA;

Query Match 99.4%; Score 165; DB 7; Length 31;
Best Local Similarity 100.0%; Pred. No. 5.4e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRTHSHMEIRTPDINPAWYASRGIRPVGR 30
Db 1 SRTHSHMEIRTPDINPAWYASRGIRPVGR 30

RESULT 13
AAW31392
ID AAW31392 standard; peptide; 32 AA.
XX
XX AAW31392;
XX
XX 06-APR-1998 (first entry)
XX
XX Human type G protein-coupled receptor ligand fragment 2.
XX
XX G protein-coupled receptor; ligand binding; pharmaceutical; modulator;
KW pituitary; central nervous system; pancreas; prophylactic;
KW therapeutic agent.
XX
XX Homo sapiens.
OS
XX WO9724436-A2.
XX
XX 10-JUL-1997.
XX
XX 26-DEC-1996; 96WO-JP003821.
XX
XX 28-DEC-1995; 95JP-00343371.
PR
XX 15-MAR-1996; 96JP-00059419.
PR
XX 12-AUG-1996; 96JP-00211805.
PR
XX 18-SEP-1996; 96JP-00246573.

XX PA (TAKE) TAKEDA CHEM IND LTD.
 XX PT Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;
 PI Kitada C;
 XX
 DR WPI; 1997-363672/33.
 DR N-PSDB; AAV02429.
 XX
 XX PT Ligand peptide for G protein-coupled receptor - acts by modulating
 PT function in the central nervous system, pancreas and pituitary gland.
 XX
 PS Claim 2; Page 185; 258pp; English.
 XX
 CC This sequence represents a peptide fragment from a novel human type
 CC ligand polypeptide corresponding to amino acid residues 23 to 54 of the
 CC sequence represented in AAW31390 and is used in an assay to monitor
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical
 CC compositions containing this ligand may be used as a pituitary function
 CC modulator, a central nervous system modulator or a pancreatic function
 CC modulator. This ligand could have specific applications as a prophylactic
 CC or therapeutic agent for dementia, depression, hyperkinetic syndrome,
 CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,
 CC growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,
 CC hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,
 CC amyotrophic lateral sclerosis, acute myocardial infarction, infertility,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis and/or oligogalactia. Assays can also be developed to screen
 CC compounds which are capable of altering the binding activity of the
 CC ligand affecting activation of the G protein-coupled receptor protein
 XX
 SQ Sequence 32 AA;

Query Match 99.4%; Score 165; DB 2; Length 32;
 Best Local Similarity 100.0%; Pred. No. 5.7e-18;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30
 Db 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30
 |||||

RESULT 14
 AAB10363
 ID AAB10363 standard; peptide; 32 AA.
 XX
 AC AAB10363;
 XX
 DT 24-NOV-2000 (first entry)
 XX
 DE Human oxytocin secretion promoting peptide SEQ ID NO: 33.
 XX
 KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;
 KW veterinary medicine; milk production.
 XX
 OS Homo sapiens.
 XX
 PN WO2000038704-A1.
 XX
 PD 06-JUL-2000.
 XX
 PF 22-DEC-1999; 99WO-JP007199.
 XX
 XX 25-DEC-1998; 98JP-00369585.
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Matsumoto H, Kitada C, Hinuma S;
 XX

DR WPI; 2000-452298/39.
 XX
 XX PT Physiologically-active polypeptide recognized as ligand by G protein-
 PT coupled receptor protein, for promoting secretion of oxytocin, as drugs
 PT for diseases relating to oxytocin secretion and in veterinary medicine.
 XX
 PS Disclosure; Page 62; 72pp; Japanese.
 XX
 CC This invention describes a novel oxytocin secretion-regulating agent
 CC which contains a ligand peptide or its salt for the G protein-coupled
 CC receptor protein. It is useful in the form of drugs for ameliorating,
 CC preventing and treating diseases relating to oxytocin secretion e.g. weak
 CC pains and atonic bleeding, before and after expulsion of placenta,
 CC uterine recovery failure, caesarean section, stoppage of artificial
 CC fertilization or galactostasis and is also applicable in veterinary
 CC medicine for promoting milk production in cow, goat and pig. This
 CC sequence represents a human peptide which acts as an oxytocin secretion
 CC promoter
 XX
 SQ Sequence 32 AA;

Query Match 99.4%; Score 165; DB 3; Length 32;
 Best Local Similarity 100.0%; Pred. No. 5.7e-18;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30
 Db 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30
 |||||

RESULT 15
 AAG62532
 ID AAG62532 standard; peptide; 32 AA.
 XX
 AC AAG62532;
 XX
 DT 24-AUG-2001 (first entry)
 XX
 DE Human CRH releasing protein related peptide SEQ ID NO: 33.
 XX
 KW Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;
 KW analgesic; hyperaldosteronism; hypercortislaemia; hypoadrenocorticism;
 KW Addison's disease; adrenal gland hyperfunction; obesity.
 XX
 OS Homo sapiens.
 XX
 PN WO200135984-A1.
 XX
 PD 25-MAY-2001.
 XX
 PF 17-NOV-2000; 2000WO-JP008119.
 XX
 PR 18-NOV-1999; 99JP-00327900.
 PR 26-SEP-2000; 2000JP-00297073.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Kitada C, Matsumoto H, Hinuma S;
 XX
 XX WPI; 2001-355552/37.
 XX
 XX PT Use of G protein receptor ligand or peptide for controlling corticotropin
 PT releasing hormone secretion.
 XX
 PS Disclosure; Page 74; 90pp; Japanese.
 XX
 CC The present sequence describes a method of controlling the secretion of
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein
 CC receptor ligand. This can be used to control the secretion of CRH and is
 CC useful as an analgesic or for treating, preventing or ameliorating
 CC diseases associated with CRH secretion such as hyperaldosteronism,
 CC hypercortislaemia, secondary or chronic hypoadrenocorticism, Addison's
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair

CC loss, and hypotension), adrenal gland hypofunction and obesity. The
CC present sequence is a peptide used in the exemplification of the
CC invention

XX

SQ Sequence 32 AA;

Query Match 99.4%; Score 165; DB 4; Length 32;
Best Local Similarity 100.0%; Pred. NO. 5.7e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHHSMEIRTPDINPAWYASRGIRPVGR 30

DB 1 SRTHHSMEIRTPDINPAWYASRGIRPVGR 30

Search completed: August 12, 2004, 14:43:54
Job time : 65.9099 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:49:10 ; Search time 54.7907 Seconds
(without alignments)
177.617 Million cell updates/sec

Title: . US-09-700-643A-2

Perfect score: 166

Sequence: 1 SRTHRHSMEIRTPDINPAWYASRGIRPVGRX 31

Scoring table: BLOSUM62

Searched: 1292805 seqs, 313927144 residues

Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_NEW_PUB.pcp.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pcp.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_NEW_PUB.pcp.*
16: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	165	99.4	31	9	US	09-933-161-15	Sequence 15, Appl
2	165	99.4	31	14	US	10-096-777-15	Sequence 15, Appl
3	165	99.4	87	13	US	10-044-932-92	Sequence 92, Appl
4	152	91.6	31	9	US	09-933-161-13	Sequence 13, Appl
5	152	91.6	31	13	US	10-044-932-39	Sequence 39, Appl
6	152	91.6	31	14	US	10-096-777-13	Sequence 13, Appl
7	152	91.6	32	13	US	10-044-932-40	Sequence 40, Appl
8	152	91.6	33	13	US	10-044-932-41	Sequence 41, Appl
9	152	91.6	98	13	US	10-044-932-28	Sequence 28, Appl
10	152	91.6	98	13	US	10-044-932-38	Sequence 38, Appl
11	152	91.6	98	13	US	10-044-932-82	Sequence 82, Appl
12	152	91.6	98	13	US	10-044-932-84	Sequence 84, Appl
13	152	91.6	98	13	US	10-044-932-86	Sequence 86, Appl
14	152	91.6	98	13	US	10-044-932-88	Sequence 88, Appl
15	147	88.6	29	13	US	10-044-932-26	Sequence 26, Appl

16	143	86.1	31	9	US-09-932-161-14	Sequence 14, Appl
17	143	86.1	31	13	US-10-044-592-4	Sequence 4, Appl
18	143	86.1	31	13	US-10-044-592-5	Sequence 5, Appl
19	143	86.1	31	13	US-10-044-592-5	Sequence 14, Appl
20	143	86.1	70	13	US-10-096-777-14	Sequence 14, Appl
21	143	86.1	70	13	US-10-044-592-90	Sequence 90, Appl
22	143	86.1	82	13	US-10-044-592-1	Sequence 1, Appl
23	143	86.1	86	13	US-10-044-592-6	Sequence 96, Appl
24	143	86.1	91	13	US-10-044-592-94	Sequence 94, Appl
25	143	86.1	25	13	US-10-044-592-78	Sequence 78, Appl
26	107	64.5	20	9	US-09-932-161-18	Sequence 18, Appl
27	107	64.5	20	14	US-10-096-777-18	Sequence 18, Appl
28	103	62.0	19	13	US-10-044-592-27	Sequence 27, Appl
29	103	62.0	20	9	US-09-932-161-16	Sequence 16, Appl
30	103	62.0	20	13	US-10-044-592-42	Sequence 42, Appl
31	103	62.0	20	14	US-10-096-777-16	Sequence 16, Appl
32	103	62.0	21	13	US-10-044-592-43	Sequence 43, Appl
33	99	59.6	22	13	US-10-044-592-44	Sequence 44, Appl
34	99	59.6	20	9	US-09-932-161-17	Sequence 17, Appl
35	99	59.6	20	13	US-10-044-592-6	Sequence 6, Appl
36	88	53.0	20	14	US-10-096-777-17	Sequence 17, Appl
37	57	34.3	9	13	US-10-044-592-80	Sequence 80, Appl
38	56.5	34.0	664	12	US-10-044-592-8	Sequence 8, Appl
39	54	32.5	209	13	US-10-389-647-469	Sequence 469, App
40	54	32.5	428	9	US-10-108-915-30	Sequence 30, Appl
41	53	31.9	133	12	US-09-820-155-4	Sequence 4, Appl
42	53	31.9	637	15	US-10-424-599-153474	Sequence 153474, A
43	52	31.3	457	12	US-10-369-493-16994	Sequence 16994, A
44	51	30.7	107	12	US-10-424-599-277367	Sequence 277367, A
45	51	30.7	465	14	US-10-424-599-228133	Sequence 228133, A
46	51	30.7	465	14	US-10-301-842-197	Sequence 197, App

ALIGNMENTS

```

RESULT 1
US-09-932-161-15
; Sequence 15, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Civeilli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; TITLE OF INVENTION: Promoting Wakefulness and Sleep
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-932-161-15

```

Query Match 99.4%; Score 165; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.2e-16;
Matches 30; Conservative 0; Mismatches 0; Indels

Qy	1	SRTHSMETPTDINPAWYASRGIRPVGR	30
Dd	1	SRTHSMETPTDINPAWYASRGIRPVGR	30

RESULT 2
US-10-096-777-15
; Sequence 15, Application US/10096777
; Publication No. US20030171270A1
; GENERAL INFORMATION:
; APPLICANT: Civalli, Olivier
; APPLICANT: Lin, Seven

; TITLE OF INVENTION: Therapeutic Compositions and Methods
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777

; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/560,915

; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15

; LENGTH: 31
; TYPE: PRT

; ORGANISM: Homo sapiens
; US-10-096-777-15

Query Match 99.4%; Score 165; DB 14; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.2e-16;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHSHMEIRTPDINPANYASRGIRPVGR 30

DB 1 SRTHSHMEIRTPDINPANYASRGIRPVGR 30

RESULT 3

; Sequence 92, Application US/10044592
; Publication No. US20020143152A1

; GENERAL INFORMATION:
; APPLICANT: Hiruma, Shuji

; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P

; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10

; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10

; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27

; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28

; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:

; SEQ ID NO 92
; LENGTH: 87

; TYPE: PRT
; ORGANISM: Homo sapiens

; US-10-044-592-92

Query Match 99.4%; Score 165; DB 13; Length 87;
Best Local Similarity 100.0%; Pred. No. 3.5e-16;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHSHMEIRTPDINPANYASRGIRPVGR 30

DB 23 SRTHSHMEIRTPDINPANYASRGIRPVGR 52

RESULT 4

; Sequence 13, Application US/09932161
; Patent No. US20020037533A1

; GENERAL INFORMATION:
; APPLICANT: Civeilli, Olivier

; TITLE OF INVENTION: Screening and Therapeutic Methods For
; FILE REFERENCE: P-UC 4679

; CURRENT APPLICATION NUMBER: US/09/932,161
; CURRENT FILING DATE: 2001-08-17

; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13

; LENGTH: 31
; TYPE: PRT

; ORGANISM: Bos taurus
; US-10-096-777-13

Query Match 91.6%; Score 152; DB 14; Length 31;
Best Local Similarity 90.0%; Pred. No. 9.2e-15;

Matches 27; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRTHSHMEIRTPDINPANYASRGIRPVGR 30

DB 1 SRAHQSHMEIRTPDINPANYAGRGIRPVGR 30

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13

; LENGTH: 31
; TYPE: PRT

; ORGANISM: Bos taurus
; US-09-932-161-13

Query Match 91.6%; Score 152; DB 9; Length 31;
Best Local Similarity 90.0%; Pred. No. 9.2e-15;

Matches 27; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRTHSHMEIRTPDINPANYASRGIRPVGR 30

DB 1 SRAHQSHMEIRTPDINPANYAGRGIRPVGR 30

RESULT 5

; Sequence 39, Application US/10044592
; Publication No. US20020143152A1

; GENERAL INFORMATION:
; APPLICANT: Hiruma, Shuji

; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P

; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10

; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10

; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27

; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28

; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:

; SEQ ID NO 39
; LENGTH: 31

; TYPE: PRT
; ORGANISM: Bovine

; US-10-044-592-39

Query Match 91.6%; Score 152; DB 13; Length 31;
Best Local Similarity 90.0%; Pred. No. 9.2e-15;

Matches 27; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRTHSHMEIRTPDINPANYASRGIRPVGR 30

DB 1 SRAHQSHMEIRTPDINPANYAGRGIRPVGR 30

RESULT 6

; Sequence 13, Application US/10096777
; Publication No. US20030171270A1

; GENERAL INFORMATION:
; APPLICANT: Civeilli, Olivier

; TITLE OF INVENTION: Therapeutic Compositions and Methods
; FILE REFERENCE: P-UC 3534

; CURRENT APPLICATION NUMBER: US/10/096,777
; CURRENT FILING DATE: 2002-03-12

; PRIOR APPLICATION NUMBER: US/09/560,915
; PRIOR FILING DATE: 2000-04-28

; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13
; LENGTH: 31

; TYPE: PRT
; ORGANISM: Bos taurus

; US-10-096-777-13

Query Match 91.6%; Score 152; DB 14; Length 31;
Best Local Similarity 90.0%; Pred. No. 9.2e-15;

Matches 27; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRTHSHMEIRTPDINPANYASRGIRPVGR 30

DB 1 SRAHQSHMEIRTPDINPANYAGRGIRPVGR 30

```
Best Local Similarity 90.0%; Pred. No. 9.2e-15; Indels 0; Gaps 0;
Matches 27; Conservative 1; Mismatches 2;

QY 1 SRTHSHMEIRTPDINPAWYASRGIRPVGR 30
||:|||||
Db 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGR 30
||:|||||

RESULT 7
US-10-044-592-40
; Sequence 40, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 40
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-40

Query Match 91.6%; Score 152; DB 13; Length 32;
Best Local Similarity 90.0%; Pred. No. 9.5e-15;
Matches 27; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRTHSHMEIRTPDINPAWYASRGIRPVGR 30
||:|||||
Db 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGR 30
||:|||||

RESULT 8
US-10-044-592-41
; Sequence 41, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 41
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-41

Query Match 91.6%; Score 152; DB 13; Length 33;
Best Local Similarity 90.0%; Pred. No. 9.8e-15;
Matches 27; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRTHSHMEIRTPDINPAWYASRGIRPVGR 30
||:|||||
Db 1 SRAHQSHMEIRTPDINPAWYAGRGIRPVGR 30
||:|||||

RESULT 9
US-10-044-592-28
; Sequence 28, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 28
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Murine
US-10-044-592-28

Query Match 91.6%; Score 152; DB 13; Length 98;
Best Local Similarity 90.0%; Pred. No. 3.1e-14;
Matches 27; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRTHSHMEIRTPDINPAWYASRGIRPVGR 30
||:|||||
Db 23 SRAHQSHMEIRTPDINPAWYAGRGIRPVGR 52
||:|||||

RESULT 10
US-10-044-592-38
; Sequence 38, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 38
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-38

Query Match 91.6%; Score 152; DB 13; Length 98;
Best Local Similarity 90.0%; Pred. No. 3.1e-14;
Matches 27; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRTHSHMEIRTPDINPAWYASRGIRPVGR 30
||:|||||
Db 23 SRAHQSHMEIRTPDINPAWYAGRGIRPVGR 52
||:|||||

RESULT 11
US-10-044-592-82
```

```
; Sequence 82, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 82
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
; US-10-044-592-82

Query Match          91.6%; Score 152; DB 13; Length 98;
Best Local Similarity 90.0%; Pred. No. 3.le-14;
Matches 27; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30
Db 23 SRAHQSMSEIRTPDINPAWYAGRGIRPVGR 52

RESULT 12
US-10-044-592-84
; Sequence 84, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 84
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
; US-10-044-592-84

Query Match          91.6%; Score 152; DB 13; Length 98;
Best Local Similarity 90.0%; Pred. No. 3.le-14;
Matches 27; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30
Db 23 SRAHQSMSEIRTPDINPAWYAGRGIRPVGR 52

RESULT 13
US-10-044-592-86
; Sequence 86, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 86
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
; US-10-044-592-86

Query Match          91.6%; Score 152; DB 13; Length 98;
Best Local Similarity 90.0%; Pred. No. 3.le-14;
Matches 27; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30
Db 23 SRAHQSMSEIRTPDINPAWYAGRGIRPVGR 52

RESULT 14
US-10-044-592-88
; Sequence 88, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 88
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
; US-10-044-592-88

Query Match          91.6%; Score 152; DB 13; Length 98;
Best Local Similarity 90.0%; Pred. No. 3.le-14;
Matches 27; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30
Db 23 SRAHQSMSEIRTPDINPAWYAGRGIRPVGR 52

RESULT 15
US-10-044-592-26
; Sequence 26, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 26
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
; US-10-044-592-26

Query Match          91.6%; Score 152; DB 13; Length 98;
Best Local Similarity 90.0%; Pred. No. 3.le-14;
Matches 27; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30
Db 23 SRAHQSMSEIRTPDINPAWYAGRGIRPVGR 52
```

```
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 26
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-26

Query Match      88.6%; Score 147; DB 13; Length 29;
Best Local Similarity 89.7%; Pred. No. 4.5e-14;
Matches 26; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 SETHRHSMEIRTPDINPAWYASRGIRPVG 29
      ||:|||||
Db      1 SRAHQHMEIRTPDINPAWYASRGIRPVG 29
      ||:|||||

Search completed: August 12, 2004, 15:22:50
Job time : 54.7907 secs
```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:36 ; Search time 17.6628 Seconds
(without alignments)
90.609 Million cell updates/sec

Title: US-09-700-643A-2

Perfect score: 166
Sequence: 1 SRRHSMETPTDINPAWYASRGIRPVGRX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
4: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	165	99.4	31	3	US-09-105-678A-9
2	165	99.4	31	3	US-09-105-678A-43
3	165	99.4	31	3	US-08-776-971-61
4	165	99.4	31	3	US-09-421-208-9
5	165	99.4	31	3	US-09-421-208-43
6	165	99.4	31	4	US-09-560-915-15
7	165	99.4	32	3	US-09-105-678A-44
8	165	99.4	32	3	US-08-776-971-62
9	165	99.4	32	3	US-09-421-208-44
10	165	99.4	33	3	US-09-105-678A-45
11	165	99.4	33	3	US-08-776-971-63
12	165	99.4	33	3	US-09-421-208-45
13	165	99.4	87	3	US-08-776-971-59
14	165	99.4	87	3	US-08-776-971-135
15	165	99.4	87	3	US-08-776-971-138
16	152	91.6	31	3	US-09-105-678A-7
17	152	91.6	31	3	US-09-105-678A-31
18	152	91.6	31	3	US-08-776-971-5
19	152	91.6	31	3	US-08-776-971-97
20	152	91.6	31	3	US-09-421-208-7
21	152	91.6	31	3	US-09-560-915-13
22	152	91.6	31	4	US-09-105-678A-32
23	152	91.6	32	3	US-08-776-971-6
24	152	91.6	32	3	US-09-421-208-32
25	152	91.6	33	3	US-09-105-678A-33
26	152	91.6	33	3	US-08-776-971-7
27	152	91.6	33	3	US-08-776-971-7

28	152	91.6	33	3	US-09-421-208-33
29	152	91.6	98	3	US-08-776-971-1
30	152	91.6	98	3	US-08-776-971-44
31	152	91.6	98	3	US-08-776-971-122
32	152	91.6	98	3	US-08-776-971-131
33	152	91.6	98	3	US-08-776-971-136
34	148	89.2	98	3	US-08-776-971-115
35	148	89.2	98	3	US-08-776-971-117
36	147	88.6	29	3	US-09-105-678A-29
37	147	88.6	29	3	US-08-776-971-3
38	147	88.6	29	3	US-09-421-208-29
39	143	86.1	31	3	US-08-776-971-8
40	143	86.1	31	3	US-03-105-678A-37
41	143	86.1	31	3	US-08-172-353-4
42	143	86.1	31	3	US-08-776-971-47
43	143	86.1	31	3	US-09-421-208-8
44	143	86.1	31	3	US-09-421-208-37
45	143	86.1	31	4	US-09-560-915-14

ALIGNMENTS

RESULT 1
US-09-105-678A-9
; Sequence 9, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09105.678A
; APPLICATION NUMBER: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/POCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-105-678A-9

Query Match 99.4%; Score 165; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 5.4e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRRHSMETPTDINPAWYASRGIRPVGR 30


```
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-421-208-9

Query Match 99.4%; Score 165; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 5.4e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30
DB 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30

RESULT 5
US-09-421-208-43
; Sequence 43, Application US/09421208
; Patent No. 6255561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400

;
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-421-208-43

Query Match 99.4%; Score 165; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 5.4e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30
DB 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30

RESULT 6
US-09-560-915-15
; Sequence 15, Application US/09560915
; Patent No. 6383764
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; TITLE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/09/560,915
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-09-560-915-15

Query Match 99.4%; Score 165; DB 4; Length 31;
Best Local Similarity 100.0%; Pred. No. 5.4e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30
DB 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30

RESULT 7
US-09-105-678A-44
; Sequence 44, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
```

; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-105-678A-44

Query Match 99.4%; Score 165; DB 3; Length 32;
Best Local Similarity 100.0%; Pred. No. 5.6e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30
Db 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30

RESULT 8
US-08-776-971-62
; Sequence 62, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; Habata, Yugo
; Kawamata, Yuji
; Hosoya, Masaki
; Fujii, Ryo
; Fukusumi, Shoji
; Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026

; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 62:
; US-08-776-971-62

Query Match 99.4%; Score 165; DB 3; Length 32;
Best Local Similarity 100.0%; Pred. No. 5.6e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30
Db 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30

RESULT 9
US-09-421-208-44
; Sequence 44, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-421-208-44

Query Match 99.4%; Score 165; DB 3; Length 32;
Best Local Similarity 100.0%; Pred. No. 5.6e-18;

```

RESULT 11
US-08-776-971-63
; Sequence 63, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; Habata, Yugo
; Kawamata, Yuji
; Hosoya, Masaki
; Fujii, Ryo
; Fukusumi, Shoji
; Kitada, Chiesko
;

```

```
Query Match          99.4%; Score 165; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 5.8e-18;
Matches 30: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30
Dp 1 SRTHRSMEIRTPDINPAWYASRGIRPVGR 30

RESULT 12
US-09-421-208-45
; Sequence 45, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA

```

; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421.208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105.678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-421-208-45

```

```

Query Match          99.4%; Score 165; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 5.8e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 SRTHSHWEIRTPDINPANYASRGIRPVGR 30
Db 1 SRTHSHWEIRTPDINPANYASRGIRPVGR 30

```

```

RESULT 13
US-08-776-971-59
; Sequence 59, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; Habata, Yugo
; Kawamata, Yuji
; Hosoya, Masaki
; Fujii, Ryo
; Fukusumi, Shoji
; Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-FEB-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371

```

```

; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 59:
; US-08-776-971-59
; Query Match          99.4%; Score 165; DB 3; Length 87;
; Best Local Similarity 100.0%; Pred. No. 1.8e-17;
; Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 SRTHSHWEIRTPDINPANYASRGIRPVGR 30
; Db 23 SRTHSHWEIRTPDINPANYASRGIRPVGR 52
;
; RESULT 14
; US-08-776-971-135
; Sequence 135, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; Habata, Yugo
; Kawamata, Yuji
; Hosoya, Masaki
; Fujii, Ryo
; Fukusumi, Shoji
; Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-FEB-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996

```

APPLICATION NUMBER: JP 8/246573
FILING DATE: 18-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 47176
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 135:
SEQUENCE CHARACTERISTICS:
LENGTH: 87 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 135:
US-08-776-971-135

Query Match 99.4%; Score 165; DB 3; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.8e-17;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHSHMEIRTPDINPAWASRGIRPVGR 30
|||||
Db 23 SRTHSHMEIRTPDINPAWASRGIRPVGR 52

RESULT 15

US-08-776-971-138
Sequence 138, Application US/08776971B
Patent No. 6228984

GENERAL INFORMATION:

APPLICANT: Hinuma, Shuji
Habata, Yugo
Kawamata, Yuji
Hosoya, Masaki
Fujii, Ryo
Fukusumi, Shoji
Kitada, Chieko
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE

NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/776,971B
FILING DATE: 06-Feb-1997
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP96/03821
FILING DATE: 28-DEC-1996
APPLICATION NUMBER: JP 7/343371
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: JP 8/59419
FILING DATE: 15-MAR-1996
APPLICATION NUMBER: JP 8/211805
FILING DATE: 12-AUG-1996
APPLICATION NUMBER: JP 8/246573
FILING DATE: 18-SEP-1996

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 47176
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 138:
SEQUENCE CHARACTERISTICS:
LENGTH: 87 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 138:
US-08-776-971-138

Query Match 99.4%; Score 165; DB 3; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.8e-17;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHSHMEIRTPDINPAWASRGIRPVGR 30
|||||
Db 23 SRTHSHMEIRTPDINPAWASRGIRPVGR 52

Search completed: August 12, 2004, 14:52:13
Job time : 17.6628 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:36 ; Search time 13.6977 Seconds
(without alignments)
217.697 Million cell updates/sec

Title: US-09-700-643A-3
Perfect score: 169
Sequence: 1 SRAHQSMETRPDINPAWYTGRIQPVGRX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*

1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	168	99.4	83	JC7607	prolactin-releasin
2	61	36.1	1236	T50904	Mg protoporphyrin
3	58	34.3	1292	T31462	probable magnesium
4	53	31.4	1798	S11120	probable untr prote
5	53	31.4	1415	T28370	conserved hypothet
6	52.5	31.1	303	2 AH2016	hypothetical prote
7	51.5	30.5	503	2 AH2193	Sun/nucleolar prot
8	51	30.2	294	2 T21075	hypothetical prote
9	50.5	29.9	176	2 S67150	hypothetical prote
10	50.5	29.9	548	2 T47548	hypothetical prote
11	50	29.6	482	1 S40887	RVS167 protein - y
12	50	29.6	962	2 H69157	excinnuclease ABC c
13	49.5	29.3	375	2 F91173	probable transport
14	49.5	29.3	375	2 F86019	probable transport
15	49.5	29.3	375	2 S47704	hypothetical 41.1K
16	49	29.0	128	2 S76955	hypothetical prote
17	49	29.0	220	2 C83292	probable glutathio
18	49	29.0	264	2 C84971	hypothetical prote
19	49	29.0	772	2 T07958	protoporphyrin IX
20	49	29.0	1193	2 T50729	magnesium-protopor
21	49	29.0	1328	2 AE2351	protoporphyrin IX
22	49	29.0	1331	2 S75000	protoporphyrin IX
23	49	29.0	1379	2 S37310	protoporphyrin IX
24	49	29.0	1380	2 S64721	protoporphyrin IX
25	49	29.0	1381	2 S71288	protoporphyrin IX
26	49	29.0	1382	2 T01789	protoporphyrin IX
27	49	29.0	1383	2 T07126	magnesium chelatase
28	48.5	28.7	106	2 H84333	hypothetical prote
29	48.5	28.7	213	2 AF0408	adenyl-lyl-sulfate k

30 48.5 28.7 430 2 P84587 hypothetical prote
31 48.5 28.7 664 2 F83376 conserved hypothet
32 48 28.4 118 2 AC3169 hypothetical prote
33 48 28.4 157 2 A81811 hypothetical prote
34 48 28.4 419 2 AH3166 hypothetical prote
35 48 28.4 455 2 D70895 probable aldC prot
36 48 28.4 498 2 T09021 beta-glucosidase h
37 48 28.4 517 2 T09022 beta-glucosidase h
38 48 28.4 719 2 S61046 ARP1 protein - yea
39 48 28.4 798 2 S29815 N-ras upstream pro
40 48 28.4 1194 2 D49851 magnesium-protopor
41 47.5 28.1 501 2 I61512 TNF receptor assoc
42 47.5 28.1 941 2 A70722 probable gcvb prot
43 47 27.8 240 2 B75318 ferritinocelin-bin
44 47 27.8 323 2 H83427 probable oxidoredu
45 47 27.8 430 1 B89009 conserved hypothet

ALIGNMENTS

RESULT 1

JC7607
prolactin-releasing peptide - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C:Accession: JC7607
R.Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Hosoya, T
Biochem. Biophys. Res. Commun. 281, 53-56, 2001
A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene: MuJ
A:Reference number: JC7607; MUID:21092785; PMID:11178959
A:Contents: Spleen
A:Accession: JC7607
A:Molecule type: DNA
A:Residues: 1-83 <YAM>
A:Cross-references: DBJ:AB040612; DDBJ:AB040613
C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior p
release, and stimulation of ACTH secretion from the pituitary.

C:Genetics:

A:Gene: PrRP

A:Introns: 33/1

Query Match 99.4%; Score 168; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 2.4e-17;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQSMETRPDINPAWYTGRIQPVGR 30

DB 22 SRAHQSMETRPDINPAWYTGRIQPVGR 51

RESULT 2

T50904
Mg protoporphyrin methyl transferase [imported] - Rubrivivax gelatinosus

C:Species: Rubrivivax gelatinosus

C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 02-Sep-2000

C:Accession: T50904

R.Nagashima, K.V.; Igarashi, N.; Harada, J.; Nagashima, S.; Matsuura, K.; Shimada, K.

submitted to the EMBL Data Library, November 1999

A:Description: Determination of Nucleotide Sequences of Rubrivivax gelatinosus Photosyn

A:Reference number: Z25270

A:Accession: T50904

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1236 <NAG>

A:Cross-references: EMBL:AB034704; PIDN:BAA94057.1

A:Experimental source: strain IL144

C:Genetics:

A:Gene: bchH

C:Superfamily: Rhodobacter capsulatus magnesium-protoporphyrin O-methyltransferase

Query Match 36.1%; Score 61; DB 2; Length 1236;

Best Local Similarity 37.5%; Pred. No. 1.7;

Matches 12; Conservative 6; Mismatches 8; Indels 6; Gaps 1;

QY 3 AHQSMETRTDPINPAWYTG-----RGIRPV 28

Db 1112 SEQVALETRTMLNPKYEGMLEHGYEGVRQI 1143

RESULT 3

T31462

probable magnesium chelatase (BC 4.99.1.-) chain H BchH - Helicobacillus mobilis

C:Species: Helicobacillus mobilis

C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 15-Sep-2000

C/Accession: T31462

R:Xiong, J.; Inoue, K.; Bauer, C.E.

Proc. Natl. Acad. Sci. U.S.A. 95, 14851-14856, 1998

A>Title: Tracking molecular evolution of photosynthesis

A/Reference number: Z21036; MUID:99061957; PMID:9843979

A/Accession: T31462

A/Status: preliminary; translated from GB/EMBL/DBS

A/Molecule type: DNA

A/Residues: 1-1292 <XIO>

A/Cross-references: EMBL:AF080002; NID:G3820536; PID:G3820560; PIDN:AAC84033.1

C/Genetics:

A/Gene: bchH

C/Superfamily: Rhodobacter capsulatus magnesium-protoporphyrin O-methyltransferase

C/Keywords: lyase

Query Match

Best Local Similarity

Matches 11; Conservative 3; Mismatches 6; Indels 6; Gaps 1;

QY 9 ETRTPDINPAWYTG-----RGIRPV 28

Db 1176 ETRTKTLNPKWEGMLKHGYEGVREI 1201

RESULT 4

S11210

probable unr protein - rat

C:Species: Rattus norvegicus (Norway rat)

C/Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999

C/Accession: S11210

R:Jeffers, M.; Paciucci, R.; Pellicer, A.

Nucleic Acids Res. 18, 4891-4899, 1990

A>Title: Characterization of unr; a gene closely linked to N-ras.

A/Reference number: S11210; MUID:90370473; PMID:2204029

A/Accession: S11210

A/Molecule type: mRNA

A/Residues: 1-798 <JEFF>

A/Cross-references: EMBL:X52311; NID:G57454; PIDN:CAA36549.1; PID:G57455

C/Keywords: DNA binding

Query Match

Best Local Similarity

Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 6 HSMETRTDPINPAWYTGIRPV 28

Db 583 HSVGITEANPTIYSGKVRPL 605

RESULT 5

C83070

conserved hypothetical protein PA4601 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Sep-2003

C/Accession: C83070

R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A/Reference number: A82950; MUID:20437337; PMID:10984043

A/Accession: C83070

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1415 <STO>

A/Cross-references: GB:AE004874; GB:AE004091; NID:G9950849; PIDN:AA07989.1; GSPDB:GN001

A/Experimental source: strain PA01

C/Genetics:

A/Gene: PA4601

C/Superfamily: oxygen sensor diguanylate cyclase/c-di-GMP phosphodiesterase

Query Match

Best Local Similarity

Matches 9; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 SRAHQSMETRTDPINPAWYTG 24

Db 330 AQARQHLLGLLPDLPQWLTDAG 353

RESULT 6

AH2016

hypothetical protein all1686 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 24-Nov-2003

C/Accession: AH2016

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anal

A/Reference number: AB1807; MUID:21595285; PMID:11759840

A/Accession: AH2016

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-303 <KUR>

A/Cross-references: GB:BA000019; PIDN:BA078052.1; PID:G17135506; GSPDB:GN00179

A/Experimental source: strain PCC 7120

C/Genetics:

A/Gene: all1686

C/Superfamily: tartrate-resistant acid phosphatase

Query Match

Best Local Similarity

Matches 15; Conservative 2; Mismatches 9; Indels 5; Gaps 2;

QY 4 HQHSMETRTDPINPAWY----TGRGIRPVGR 30

Db 226 HERSYE-RTRADTGYTLTCGAGAGNRPVGR 255

RESULT 7

A82193

Sun/nucleolar protein family protein VC1502 [imported] - Vibrio cholerae (strain N16961)

C:Species: Vibrio cholerae

C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C/Accession: A82193

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A>Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A/Reference number: A82035; MUID:20406833; PMID:10952301

A/Accession: A82193

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-503 <HEI>

A/Cross-references: GB:AE004228; GB:AE003852; NID:G96555997; PIDN:AAF94657.1; GSPDB:GN001

A/Experimental source: serogroup O1; strain N16961; biotype El Tor

C/Genetics:

A/Gene: VC1502

A/Map position: 1

Query Match

Score 30.5%; DB 2; Length 503;

DB 120 SECHQHNVFYFLPAVDLTKQWFIAGFEQVG 151

RESULT 10
T47548
hypothetical protein F8J2.80 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47548
R;Nyakatura, G.; Fartmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichselgartner, M.;
Mayer, K.F.X.
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24458
A:Accession: T47548
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-548 <MYA>
A:Cross-references: EMBL:AL132969
A:Experimental source: cultivar Columbia; BAC clone F8J2
C:Genetics:
A:Map position: 3
A:Introns: 78/1; 143/1; 242/2; 377/3; 428/2; 447/3; 470/3; 502/3; 532/3
A:Note: F8J2.80

Query Match 29.9%; Score 50.5; DB 2; Length 548;
Best Local Similarity 35.5%; Pred. No. 24;
Matches 11; Conservative 3; Mismatches 10; Indels 7; Gaps 1;

QY 1 SRAHQ-----HSMETRTDPINPAWYTGKG 24
DB 118 SLHQSPMFLHHPHQSHFPHHQPSWYWGK 148

RESULT 11
S40887
RVS167 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YDR388W
C:Species: Saccharomyces cerevisiae
C:Date: 31-Mar-1992 #sequence_revision 06-Feb-1995 #text_change 21-Jul-2000
C:Accession: S40887; S69672
R;Bauer, F.; Urdaci, M.; Aigle, M.; Crouzet, M.
Mol. Cell. Biol. 13, 5070-5084, 1993
A:Title: Alteration of a Yeast SH3 protein leads to conditional viability with defects
A:Reference number: S40887; MUID:93330299; PMID:8336735
A:Accession: S40887
A:Molecule type: DNA
A:Residues: 1-482 <BAU>
A:Cross-references: EMBL:M92092; NID:9172615; PIDN:AAA35051.1; PID:9172616
R;Dietrich, F.S.
submitted to the EMBL Data Library, July 1995
A:Description: The sequence of S. cerevisiae cosmid 9481, 9509, 9926, 9461, and lambda
A:Reference number: S6965
A:Accession: S69672
A:Molecule type: DNA
A:Residues: 1-482 <DIE>
A:Cross-references: EMBL:U32274; NID:927313; PIDN:AAB64830.1; PID:9273321; GSPDB:GN000
C:Genetics:
A:Gene: SGD:RVS167; MIPS:YDR388W
A:Cross-references: SGD:50002796; MIPS:YDR388W
A:Map position: 4R
C:Superfamily: RVS167 protein; RVS161 protein homology; SH3 homology
C:Keywords: transmembrane protein
F:4-270/Domain: RVS161 protein homology <RVS>
F:292-422/Region: alanine/glycine/proline-rich
F:428-477/Domain: SH3 homology <SH3>

	Best Local Similarity	50.0%; Pred. No. 25;				
	Matches 11;	Conservative 3;	Mismatches	2;	Indels	6; Gaps 2;
Qy	11	RTPDINPAWYTG	----GIRP	27		
		:		:	:	:
Db	453	RTPDVN-EWWTGRY	QQGVFP	473		

RESULT 12 H69157

excinuclease ABC chain A - Methanobacterium thermoautotrophicum (strain Delta H)
N:Alternate names: uvrA protein
N:Contains: excision endonuclease ABC (EC 3.1.1.-.-) chain A
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 02-Feb-2001
C:Accession: H69157
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; diwani, N.; Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J: Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funcu
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: H69157
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-962 <MTH>
A:Cross-references: GB:AE000828; GB:AE000666; MID:G2621504; PIDN:AAB84949.1; PID:G262150
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH443

A:Start codon: TTG
C:Superfamily: excinuclease ABC chain A; ATP-binding cassette homology
C:Keywords: ATP; DNA binding; DNA repair; duplication; hydrolase; nucleotide binding; P-
F:38-45/Region: nucleotide-binding motif A (P-loop)
F:632-915/Domain: ATP-binding cassette homology <ABC>
F:649-656/Region: nucleotide-binding motif A (P-loop)

Query Match 29.6%; Score 50; DB 2; Length 962;
Best Local Similarity 40.5%; Pred. No. 53;

Matches 15; Conservative 0; Mismatches 4; Indels 18; Gaps 2;

QY 11 RTPDINPAWYTG-----RGTRPVR 30

DB 703 RTPSRPATVGTGVTTHRELFQTPKARKGYRP-GR 738

RESULT 13

Probable transporter ECs4358 [imported] - Escherichia coli (strain O157:H7, substrain R1
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: F91173

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
R:Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: F91173

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-375 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA037781.1; PID:gl3363832; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain R1MD 0509952

C:Genetics:
A:Gene: ECs4358

Query Match 29.3%; Score 49.5; DB 2; Length 375;
Best Local Similarity 47.4%; Pred. No. 22;

Matches 9; Conservative 5; Mismatches 2; Indels 3; Gaps 1;

QY 7 SMETR---TPDINPAWYTG 22

DB 161 SLETRMRFPNLDPAWFGG 179

RESULT 14

F86019
probable transporter yhhJ [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: F86019

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouisis, K.; Apodaca,
Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: F86019

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-375 <STO>

A:Cross-references: GB:AE005174; MID:gl2518177; PIDN:AAG58618.1; GSPDB:GN00145; UWGP:248

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:
A:Gene: yhhJ

Query Match 29.3%; Score 49.5; DB 2; Length 375;
Best Local Similarity 47.4%; Pred. No. 22;

Matches 9; Conservative 5; Mismatches 2; Indels 3; Gaps 1;

QY 7 SMETR---TPDINPAWYTG 22

DB 161 SLETRMRFPNLDPAWFGG 179

RESULT 15

Hypothetical 41.1K protein (rhaB-pit intergenic region) - Escherichia coli (strain K-12)
N:Alternate names: yhhJ protein
C:Species: Escherichia coli
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 01-Mar-2002

C:Accession: S47704; H65145

R:Plunkett, G.

submitted to the EMBL Data Library, March 1994

A:Reference number: S47666

A:Accession: S47704

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-375 <PLU>

A:Cross-references: EMBL:U00039; MID:G466582; PIDN:AAB18460.1; PID:G466621

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: H65145

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-375 <BLAT>

A:Cross-references: GB:AE000424; GB:U00096; MID:G2367230; PIDN:AAC76510.1; PID:gl789897;

A:Experimental source: strain K-12, substrain MG1655

C:Genetics:
A:Gene: yhhJ

Query Match 29.3%; Score 49.5; DB 2; Length 375;
Best Local Similarity 47.4%; Pred. No. 22;

Matches 9; Conservative 5; Mismatches 2; Indels 3; Gaps 1;

QY 7 SMETR---TPDINPAWYTG 22

DB 161 SLETRMRFPNLDPAWFGG 179

Search completed: August 12, 2004, 14:50:27

Job time: 13.6977 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 8.11047 Seconds
(without alignments)
199.024 Million cell updates/sec

Title: US-09-700-643A-3

Perfect score: 169

Sequence: 1 SRAHQSMETRTDPINPAWYTGRIPIVGRX 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	168	99.4	83	1 PRRP_RAT	P81278 rattus norv
2	157	92.9	98	1 PRRP_BOVIN	P81264 bos taurus
3	143	84.6	87	1 PRRP_HUMAN	P81277 homo sapien
4	56	33.1	428	1 NER3_BOVIN	O97859 bos taurus
5	53	31.4	798	1 UNR_EAT	P18395 rattus norv
6	51	30.2	428	1 NER3_HUMAN	Q9uq49 homo sapien
7	50.5	29.9	772	1 LM8T_HUMAN	Q9y468 homo sapien
8	50	29.6	482	1 R167_YEAST	P39743 saccharomyc
9	50	29.6	962	1 UVR4_METTH	O26543 methanobact
10	49.5	29.3	374	1 YHJH_ECOLI	P31993 escherichia
11	49	29.0	264	1 Y35E_BUCAL	P57436 buchnera ap
12	49	29.0	1193	1 ECHI_RHOSH	Q9rfd5 rhodobacter
13	48.5	28.7	213	1 CYSC_YERPE	Q82bp3 yersinia pe
14	48	28.4	413	1 EX7L_COREF	Q8fgp1 corynebacte
15	48	28.4	417	1 EX7L_COREL	Q8nmr3 corynebacte
16	48	28.4	719	1 NRPI_YEAST	P32770 saccharomyc
17	48	28.4	798	1 UNR_HUMAN	O75534 homo sapien
18	48	28.4	1087	1 XPO7_HUMAN	Q9uia9 homo sapien
19	48	28.4	1087	1 XPO7_MOUSE	Q9epk7 mus musculu
20	48	28.4	1194	1 ECHI_RHOCA	P26162 rhodobacter
21	47.5	28.1	501	1 TR2A_MOUSE	P39429 mus musculu
22	47.5	28.1	941	1 GCSP_MYCTU	Q50601 mycobacteri
23	47	27.8	453	1 TBB2_GEOCN	P32925 geotrichum
24	47	27.8	940	1 UVR4_VIBPA	Q871a0 vibrio para
25	47	27.8	940	1 UVR4_VIBVU	Q8dcj3 vibrio vuln
26	47	27.8	1083	1 T2D3_HUMAN	O00268 homo sapien
27	46.5	27.5	272	1 TRA2_DROVI	O02008 drosophila
28	46.5	27.5	652	1 TETP_CLOPE	Q46306 clostridium
29	46.5	27.5	970	1 UVR4_SYNY3	P73412 synechocyst
30	46	27.2	365	1 QUEA_RICCN	Q921v7 rickettsia
31	46	27.2	569	1 UVR4_VITST	O08516 vitreoscill
32	46	27.2	598	1 RF2P_DROSI	Q24629 drosophila
33	46	27.2	955	1 UVR4_ZYMMO	O31151 zymomonas m

ALIGNMENTS

RESULT 1

ID	PRRP_RAT	STANDARD	PRT	83 AA
DT	P81278; Q8K3Y0;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing			
DE	hormone) [Contains: Prolactin-releasing peptide PrRP1; Prolactin-			
DE	releasing peptide PrRP20].			
GN	PRH.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_taxid=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RC	TISSUE=Brain;			
RX	MEDLINE=98268781; PubMed=9607765;			
RA	Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,			
RA	Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,			
RA	Kurokawa T., Nishimura O., Onda H., Fujino M.,			
RT	"A prolactin-releasing peptide in the brain.";			
RL	Nature 393:272-276(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RC	STRAIN=Sprague-Dawley; TISSUE=Hypothalamus;			
RA	Anderson S.T., Kokay I.C., Lang T., Grattan D.R., Curlewis J.D.;			
RT	"Quantitation of prolactin-releasing peptide (PrRP) mRNA expression in			
RT	specific brain regions during the rat oestrous cycle and in			
RT	lactation.";			
RL	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	TISSUE SPECIFICITY.			
RX	MEDLINE=99426652; PubMed=10498338;			
RA	Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,			
RA	Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,			
RA	Sumino Y., Fujino M.;			
RT	"Tissue distribution of prolactin-releasing peptide (PrRP) and its			
RT	receptor.";			
RL	Regul. Pept. 83:11-10(1999).			
CC	-!- FUNCTION: Stimulates prolactin (PRL) release and regulates the			
CC	expression of prolactin through its receptor GPR10. May stimulate			
CC	lactotrophs directly to secrete PRL.			
CC	-!- ALTERNATIVE PRODUCTS:			
CC	Event=Alternative splicing; Named isoforms=2;			
CC	Name=1;			
CC	Isoid=P81278-1; Sequence=Displayed;			
CC	Name=2;			
CC	Isoid=P81278-2; Sequence=VSP_004370;			
CC	-!- TISSUE SPECIFICITY: Widely expressed, with highest levels in			
CC	medulla oblongata and hypothalamus.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			

Q8X5u9 escherichia
Q8fb02 escherichia
P07671 escherichia
Q9kuw5 vibrio chol
P37434 salmonella
Q8X5z0 streptococc
Q8zj2 streptococc
Q9zy84 streptococc
P4410 haemophilus
P57979 pasteurilla
P72481 streptococc
Q97sx7 streptococc

QY 1 SRAHSHMETRTPDINPAWYTGIRPVR 30
DB 23 SRTHSHMETRTPDINPAWYASRGIRPVR 52

RESULT 4
NER3_BOVIN
ID_NER3_BOVIN STANDARD; PRT; 428 AA.
AC C97859;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sialidase 3 (EC 3.2.1.18) (Membrane sialidase)
DE (N-acetyl-alpha-neuraminidase 3).
GN NEU3.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.
RC TISSUE=Brain;
RX MEDLINE=99143165; PubMed=9988745;
RA Miyagi T., Wada T., Iwamatsu A., Hata K., Yoshikawa Y., Tokuyama S.,
SAWADA M.;
RT "Molecular cloning and characterization of a plasma membrane-
associated sialidase specific for gangliosides.";
RL J. Biol. Chem. 274:5004-5011(1999).
CC -!- FUNCTION: Plays a role in modulating the ganglioside content of
the lipid bilayer at the level of membrane-bound sialyl
glycoconjugates.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2-3)-, alpha-(2-6)-,
alpha-(2-8)-glycosidic linkages of terminal sialic residues in
oligosaccharides, glycoproteins, glycolipids, colominic acid and
synthetic substrates.
CC -!- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed in brain.
CC -!- SIMILARITY: Belongs to family 33 of glycosyl hydrolases.
CC -!- SIMILARITY: Contains 3 BNR repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; A5008184; BAA75071.1; -.
DR InterPro; IPR002860; GH_BNR.
DR Pfam; PF02012; BNR; 3.
KW Hydrolase; Glycosidase; Membrane; Repeat.
FT REPEAT 129 140
FT REPEAT 203 214 BNR 1.
FT REPEAT 254 265 BNR 2.
FT SITE 24 27 BNR 3.
FT ACT_SITE 25 25 By similarity.
FT ACT_SITE 45 45 Potential.
FT ACT_SITE 50 50 Potential.
FT ACT_SITE 87 87 Potential.
FT ACT_SITE 225 225 Potential.
FT ACT_SITE 245 245 Potential.
FT ACT_SITE 341 341 By similarity.
FT ACT_SITE 371 371 Potential.
FT ACT_SITE 388 388 Potential.
SQ SEQUENCE 428 AA; 47916 MW; 418B34F3245A8F21 CRC64;

Query Match 33.1%; Score 56; DB 1; Length 428;
Best Local Similarity 37.0%; Pred. No. 1.3;
Matches 10; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 2 RAHQSHMETRTPDINPAWYTGIRPV 28
DB 195 RARPHSLMIYSDDLGATWHGRLIKPM 221

RESULT 5
UNR_RAT
ID_UNR_RAT STANDARD; PRT; 798 AA.
AC P18395;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE UNR protein.
GN UNR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=90378473; PubMed=2204029;
RA Jeffers M., Paciucci R., Pellicer A.;
RT "Characterization of unr; a gene closely linked to N-ras.";
RL Nucleic Acids Res. 18:4891-4899(1990).
CC -!- FUNCTION: RNA-binding protein (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the cold-shock domain (CSD) family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X52311; CAA36549.1; -.
DR PIR; S11210; S11210.
DR HSP; P15277; 1MJC.
DR InterPro; IPR002059; Cold shock.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR Pfam; PF00313; CSD; 7.
DR ProDom; PDC00621; Cold_shock; 1.
DR SMART; SM00357; CSP; 5.
DR PROSITE; PS00352; COLD_SHOCK; 4.
KW RNA-binding; Repeat.
FT DOMAIN 26 87 CSD 1.
FT DOMAIN 136 179 CSD 2 (INCOMPLETE).
FT DOMAIN 186 245 CSD 3.
FT DOMAIN 297 337 CSD 4 (INCOMPLETE).
FT DOMAIN 349 410 CSD 5.
FT DOMAIN 447 507 CSD 6.
FT DOMAIN 519 579 CSD 7.
FT DOMAIN 610 670 CSD 8.
FT DOMAIN 674 735 CSD 9.
SQ SEQUENCE 798 AA; 88894 MW; F484B3FA8B0995A4 CRC64;

Query Match 31.4%; Score 53; DB 1; Length 798;
Best Local Similarity 43.5%; Pred. No. 6.8;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 6 HSMETRTPDINPAWYTGIRPV 28
DB 593 HSVNGITEEANTPTIYSGKVRPL 605

RESULT 6
NER3_HUMAN
ID_NER3_HUMAN STANDARD; PRT; 428 AA.
AC Q9UQ49; Q9NQ51;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sialidase 3 (EC 3.2.1.18) (Membrane sialidase) (Ganglioside sialidase)
DE (N-acetyl-alpha-neuraminidase 3).
GN NEU3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=9933533; PubMed=10405317;
RA Wada T., Yoshikawa Y., Tokuyama S., Kuwabara M., Akita H., Miyagi T.;
RT "Cloning, expression, and chromosomal mapping of a human ganglioside
RL sialidase.";
RN Biochem. Biophys. Res. Commun. 261:21-27(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX PubMed=10861246;
RA Monti E., Bassi M.T., Papini N., Riboni M., Manzoni M., Venerando B.,
RT Croci G., Preti A., Ballabio A., Tettamanti G., Borsani G.;
RT "Identification and expression of NEU3, a novel human sialidase
RL associated to the plasma membrane.";
RN Biochem. J. 349:343-351(2000).
CC -1- FUNCTION: Plays a role in modulating the ganglioside content of
CC the lipid bilayer at the level of membrane-bound sialyl
CC glycoconjugates.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2-3)-, alpha-(2-6)-,
CC alpha-(2-8)-glycosidic linkages of terminal sialic residues in
CC oligosaccharides, glycoproteins, glycolipids, colominic acid and
CC synthetic substrates.
CC -1- SUBCELLULAR LOCATION: Membrane-associated.
CC -1- TISSUE SPECIFICITY: Highly expressed in skeletal muscle, testis,
CC adrenal gland and thymus, followed by pancreas, liver, heart and
CC thymus. Weakly expressed in kidney, placenta, brain and lung.
CC -1- MISCELLANEOUS: Optimum pH is 3.8.
CC -1- SIMILARITY: Belongs to family 33 of glycosyl hydrolases.
CC -1- SIMILARITY: Contains 3 BNR repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB008185; BAA82611.1; -;
DR EMBL; Y18563; CAB96131.1; ALT_INIT.
DR Genbank; HGNC:7760; NEU3.
DR MIM; 604617; -;
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0003824; F:catalytic activity; TAS.
DR GO; GO:0006689; P:ganglioside catabolism; TAS.
DR InterPro; IPR002860; GH_ENR.
DR Pfam; PF02012; BNR; 3
KW Hydrolase; Glycosidase; Membrane; Repeat.
FT REPEAT 129 140 BNR 1.
FT REPEAT 203 214 BNR 2.
FT REPEAT 254 265 BNR 3.
FT SITE 24 27 FRIP MOTIF.
FT ACT_SITE 25 25 By similarity.
FT ACT_SITE 45 45 Potential.
FT ACT_SITE 50 50 Potential.
FT ACT_SITE 87 87 Potential.
FT ACT_SITE 225 225 Potential.
FT ACT_SITE 245 245 Potential.
FT ACT_SITE 340 340 By similarity.
FT ACT_SITE 370 370 Potential.
FT ACT_SITE 387 387 Potential.
SQ SEQUENCE 428 AA; 48252 MW; 35D1DD9359A78C98 CRC64;

Query Match 30.2%; Score 51; DB 1; Length 428;
Best Local Similarity 33.3%; Pred. No. 6.9;
Matches 9; Conservative 6; Mismatches 12; Indels 0; Gaps 0;
Oy 2 RAHQHSMETRTDIPINFAWYTGCRPV 28
Db 195 KTRPSLMYSDDLGVTHHGRILRM 221
RESULT 7
LMET_HUMAN STANDARD; PRT; 772 AA.
AC Q9V468; Q9H126; Q9H1G5; Q9UG06; Q9UJB9; Q9Y4C9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Lethal(3)malignant brain tumor-like protein (L(3)mbl-like) (L(3)mbl
DE protein homolog) (H-1(3)mbl protein) (H-L(3)MBT).
GN L3MBTL OR L3MBT OR KIAA0881.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Brain;
RX MEDLINE=99373015; PubMed=10445843;
RA Koga H., Matsui S.-I., Hirota T., Takebayashi S.-I., Okumura K.,
RT Saya H.;
RT "A human homolog of Drosophila lethal(3)malignant brain tumor
RL (L(3)mbl) protein associates with condensed mitotic chromosomes.";
RN Oncogene 18:3799-3809(1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RC TISSUE=Uterus;
RA Koshier K., Beyer A., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Coley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leiva-Salao M.H., Levensha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McDonach L.J., McLeay K., Nickerson T.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Peck A.I.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Symamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Williams S.A.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN [4]
RP SEQUENCE OF 215-772 FROM N.A. (ISOFORM 4).
RC TISSUE=Brain;
RX MEDLINE=98403880; PubMed=9734811;
RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,
RA Kotani H., Nomura N., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. X.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro.";
 RL DNA Res. 5:169-176(1998).
 RN [5]
 RP INTERACTION WITH ETV6.
 RX MEDLINE=22590457; PubMed=12588862;
 RA Bocconi P., MacGrogan D., Scandura J.M., Nimer S.D.;
 RA "The human L(3)MBT Polycomb group protein is a transcriptional
 RT repressor and interacts physically and functionally with TEL
 RT (ETV6).";
 RL J. Biol. Chem. 278:15412-15420(2003).
 CC -!- FUNCTION: Polycomb group (PcG) protein. PcG proteins maintain the
 CC transcriptionally repressive state of genes, probably via a
 CC modification of chromatin, rendering it heritably changed in its
 CC expressibility. Participates to the ETV6-mediated repression.
 CC Probably plays a role in cell proliferation. Overexpression
 CC induces multinucleated cells, suggesting that it is required to
 CC accomplish normal mitosis.
 CC -!- SUBUNIT: Homodimer. Interacts with ETV6.
 CC -!- SUBCELLULAR LOCATION: Nuclear; excluded from the nucleolus. Does
 CC not colocalizes with the PcG protein BMI1, suggesting that these
 CC two proteins do not belong to the same complex.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=4;
 CC Name=1; Synonyms=mbt-I;
 CC IsoId=Q9Y468-1; Sequence=Displayed;
 CC Name=2; Synonyms=mbt-II;
 CC IsoId=Q9Y468-2; Sequence=VSP_003902;
 CC Name=3;
 CC IsoId=Q9Y468-3; Sequence=VSP_003901, VSP_003902;
 CC Name=4;
 CC IsoId=Q9Y468-4; Sequence=VSP_003903;
 CC -!- TISSUE SPECIFICITY: Widely expressed. Expression is reduced in
 CC colorectal cancer cell line SW480 and promyelocytic leukemia cell
 CC line HL-60.
 CC -!- DEVELOPMENTAL STAGE: In interphase cells, it is scattered
 CC throughout the nucleoplasm. In mitotic cells, it strongly
 CC associates with condensed chromosomes from the prophase to
 CC telophase.
 CC -!- SIMILARITY: Contains 3 mbt domains.
 CC -!- CAUTION: Ref.3 (CAC18508) sequences differ from that shown due to
 CC erroneous gene model prediction.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U99358; AAC59438.1; -;
 DR EMBL; AL110279; CAB53714.1; -;
 DR EMBL; Z98752; CAC16799.1; -;
 DR EMBL; Z98752; CAC16800.1; -;
 DR EMBL; Z98752; CAC18508.1; -;
 DR EMBL; AL031681; CAC17518.1; ALT_SEQ.
 DR EMBL; AL031681; CAB43959.1; -;
 DR EMBL; AB014581; BAA31656.1; -;
 DR F1R; T14794; T14794.
 DR Genew; HGNC:15905; L3MBTL.
 DR GK; Q9Y468; -;
 DR InterPro; IPR004092; Mbt.
 DR InterPro; IPR002515; Znf_C2HC.
 DR Pfam; PF02820; mbt; 3.
 DR Pfam; PF01530; zf-C2HC; 1.
 DR SMART; SM00561; Mbt; 3.
 DR Transcription regulation; Repressor; Chromatin regulator; Zinc-finger;
 KW DNA-binding; Nuclear protein; Repeat; Alternative splicing.
 FT DOMAIN 242 315
 FT MFT 1.
 FT MFT 2.
 FT MFT 3.
 FT DOMAIN 349 422
 FT MFT 1.
 FT MFT 2.
 FT MFT 3.
 FT DOMAIN 453 526
 FT MFT 1.
 FT MFT 2.
 FT MFT 3.

FT ZN FING 552 578 C2HC-TYPE
 FT VARSPLIC 1 348 Missing (in isoform 3).
 FT /FTid=VSP_003901.
 FT ARIVRVTHVSKTLVWTVAQLGDLVCSHDLQEGKILETGV
 FT HSLICSLPTHLAKLSFASDSQY -> VRCKRVGDRAGVT
 FT VLKTAGSRCPPQRHFC (in isoform 2 and
 FT isoform 3).
 FT /FTid=VSP_003902.
 FT ARIVRVTHVSKTLVWTVAQLGDLVCSHDLQEGKILETGV
 FT HSLICSLPTHLAKLSFASDSQY -> MIDGEAFLLLTQAD
 FT IVKIMSVKLPALIKYINAILMFKNADDTLK (in
 FT isoform 4).
 FT /FTid=VSP_003903.
 FT P -> L (IN REF. 1).
 FT CONFLICT 305 305
 FT CONFLICT 320 321 LR -> MC (IN REF. 1).
 FT CONFLICT 332 332 L -> M (IN REF. 1).
 FT CONFLICT 595 595 S -> P (IN REF. 1).
 SQ SEQUENCE 772 AA; 85916 MW; 117B03A628826B29 CRC64;
 Query Match 29.9%; Score 50.5; DB 1; Length 772;
 Best Local Similarity 36.7%; Pred. No. 15;
 Matches 11; Conservative 6; Mismatches 10; Indels 3; Gaps 2;
 QY 1 SRAHQHSMETPTDINPA-WY--TGRGIRP 27
 DB 276 SECHDFWYNANSPDIHPAGWFEKTHKIQP 305
 RESULT 8
 RI67_YEAST
 ID RI67_YEAST STANDARD; PRT; 482 AA.
 AC P39743;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Reduced viability upon starvation protein 167.
 GN RVSL167 OR YDR388W OR D9509.8.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=X2180;
 RX MEDLINE=93330299; PubMed=8336735;
 RA Bauer F., Urdaci M., Aigle M., Crouzet M.;
 RT "Alteration of a yeast SH3 protein leads to conditional viability
 RT with defects in cytoskeletal and budding patterns.";
 RL Mol. Cell. Biol. 13:5070-5084(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97313263; PubMed=9169867;
 RA Jacq C., Alt-Moerbe J., Andre B., Arnold W., Bahr A., Ballesta J.P.G.,
 RA Baigues M., Baron L., Becker A., Biteau N., Bloeker H., Blugeon C.,
 RA Boskovic J., Brandt P., Brueckner M., Buitrago M.J., Coster F.,
 RA Delaveau T., del Rey F., Dujon B., Eide L.G., Garcia-Cantalejo J.M.,
 RA Goffeau A., Gomez-Peris A., Granotier C., Hanemann V., Hankeln T.,
 RA Hoesel J.D., Jaeger W., Jimenez A., Joniaux J.-L., Kraemer C.,
 RA Kuester H., Laamanen P., Legros Y., Louis E.J., Moeller-Rieker S.,
 RA Monnet A., Moro M., Mueller-Auer S., Nussbaumer B., Paricio N.,
 RA Paulin L., Perea J., Perez-Alonso S., Perez-Ortin J.E., Pohl T.M.,
 RA Prydz H., Purnelle B., Rasmussen S.W., Remacha M., Revuelta J.L.,
 RA Rieger M., Salom D., Saluz H.P., Saiz J.E., Saren A.-M., Schaefer M.,
 RA Scharfe M., Schmidt E.R., Schneider C., Scholler P., Schwarz S.,
 RA Urrestarazu L.A., Verhasselt P., Vissers S., Voet M., Volckaert G.,
 RA Wagner G., Wambutt R., Wedler E., Wedler H., Woelfl S., Harris D.B.,
 RA Bowman S., Brown D., Churcher C.M., Connor R., Dedman K., Gentles S.,
 RA Hamlin N., Hunt S., Jones L., McDonald S., Murphy L., Niblett D.,
 RA Odell C., Oliver K., Rajandream M.A., Richards C., Shore L.,
 RA Walsh S.V., Barrell B.G., Dietrich F.S., Mulligan J.T., Allen E.,
 RA Araujo R., Aviles E., Berio A., Carpenter J., Chen E., Cherry J.M.,
 RA Lushkari D., Duncan M., Hunicke-Smith S., Hyman R.W., Komp C.,
 RA Lashkari D., Lew H., Lin D., Mosedale D., Nakahara K., Namath A.,

RA Oefner P., Oh C., Petel F.X., Roberts D., Schramm S., Schroeder M.,
RA Shogren T., Shroff N., Winant A., Yelton M.A., Botstein D.,
RA Davis R.W., Johnston M., Andrews S., Brinkman R., Cooper J., Ding H.,
RA Du Z., Favell A., Fulton L., Gattung S., Greco T., Hallsworth K.,
RA Hawkins J., Hillier L.W., Jier M., Johnson D., Johnson L.,
RA Kirsten J., Kucaba T., Langston Y., Latreille P., Le T., Mardis E.,
RA Meneses S., Miller N., Nhan M., Pauley A., Peluso D., Rifkin L.,
RA Riles L., Raich A., Trevisan E., Vignati D., Wilcox L., Wohlman P.,
RA Vaudin M., Wilson R., Waterston R., Albermann K., Hanl J., Heumann K.,
RA Kline K., Mewes H.-W., Zollner A., Zaccaria P.;
RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome IV.";
RL Nature 387:75-78(1997).
RN [3]
RP ACTIN-BINDING.
RX MEDLINE=95236199; PubMed=7719850;
RA Amberg D.C., Basart E., Botstein D.;
RT "Defining protein interactions with yeast actin in vivo.";
RL Nat. Struct. Biol. 2:28-35(1995).
CC -!- FUNCTION: Component of a cytoskeletal structure that is required
CC for the formation of endocytic vesicles at the plasma membrane
CC level. Could be implicated in cytoskeletal reorganization in
CC response to environmental stresses and could act in the budding
CC site selection mechanism. Binds to actin.
CC -!- SIMILARITY: Contains 1 BAR domain.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M92092; AAA35051.1; -;
DR EMBL; U32274; AAB64830.1; -;
DR PIR; S40887; S40887.
DR HSP; P19174; 2HSP.
DR GERMOnline; 140880; -;
DR SGD; S0002796; RV5167.
DR GO; GO:0005857; C:actin cortical patch (sensu Saccharomyces); IDA.
DR GO; GO:0008092; F:cytoskeletal protein binding; IPI.
DR GO; GO:0006897; P:endocytosis; IMP.
DR GO; GO:0007121; P:polar budding; IMP.
DR GO; GO:0006570; P:response to osmotic stress; IMP.
DR InterPro; IPR006632; BAR.
DR InterPro; IPR004148; BAR_dom.
DR InterPro; IPR001452; SH3.
DR Pfam; PF03114; BAR; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRODOM; PD000066; SH3; 1.
DR SMART; SM00721; BAR; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS0002; SH3; 1.
KW Cytoskeleton; SH3 domain; Coiled coil; Transmembrane; Actin-binding.
FT DOMAIN 31 64
FT COILED COIL (POTENTIAL).
FT DOMAIN 174 204
FT COILED COIL (POTENTIAL).
FT DOMAIN 292 344
FT ALA/GLY/PRO-RICH.
FT TRANSMEM 344 367
FT POTENTIAL.
FT DOMAIN 421 482
FT SH3.
SQ SEQUENCE 482 AA; 52774 MW; 3F0AB53BEC95A5B CRC64;
Query Match 29.6%; Score 50; DB 1; Length 482;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 11; Conservative 3; Mismatches 2; Indels 6; Gaps 2;
QY 11 RTPDINPAWYTGK-----GIRP 27
Db 453 RTPDIN-EWWTGRNGGQGVFP 473
RESULT 9

UVRA METHTH STANDARD; PRT; 962 AA.
ID UVRA METHTH
AC O6543; 2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DT UVABC system protein A (UvrA protein) (Excinuclease ABC subunit A).
DE UVRA OR MTH443.
GN Methanobacterium thermoautotrophicum.
OS Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzaden R., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Harrison D., Hoang L., Wang Y., Wierzbowski J., Gibson R.,
RA Spadafora R., Vicare R., Wang H., Safer H., Patwell D., Prabhakar S.,
RA Jiwan N., Caruso A., Bush D., Goval A., Pietrowski S., Church G.M.,
RA McDougall S., Shimer G., Rice P., Noelling J., Reeve J.N.;
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT delat: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -!- FUNCTION: The UvrABC repair system catalyzes the recognition and
CC processing of DNA lesions. UvrA is an ATPase and a DNA-binding
CC protein. A damage recognition complex composed of 2 uvrA and 2
CC uvrB subunits scans DNA for abnormalities. When the presence of a
CC lesion has been verified by uvrB, the uvrA molecules dissociate
CC (By similarity).
CC -!- SUBUNIT: Forms a heterotrimer with uvrB during the search for
CC lesions (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the ABC transporter family. UvrA subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000828; AAB84949.1; -;
DR PIR; H69157; H69157.
DR HAMAP; MF_00205; -; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR004602; UvrA.
DR Pfam; PF00005; ABC_tran; 2.
DR PRODOM; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 2.
DR TIGRfams; TIGR00630; uvrA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW SOS response; Excision nuclease; DNA repair; DNA recombination;
KW DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;
KW Zinc-finger; Complete proteome.
FT NP_BIND 38 45
FT ATP (POTENTIAL).
FT NP_BIND 649 656
FT ATP (POTENTIAL).
FT ZN_FING 748 774
FT C4-TYPE.
SQ SEQUENCE 962 AA; 108395 MW; 2C0EF7FC41CCD060 CRC64;
Query Match 29.6%; Score 50; DB 1; Length 962;
Best Local Similarity 40.5%; Pred. No. 23;
Matches 15; Conservative 0; Mismatches 4; Indels 18; Gaps 2;
QY 11 RTPDINPAWYTGK-----RGIRPVR 30
Db 703 RTPSRNPATYGVTHIRLFAQTPEARKGGRP-GR 738


```
RESULT 10
YHHJ_ECOLI
ID YHHJ_ECOLI STANDARD; PRT; 374 AA.
AC P31993; P76703;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Hypothetical protein yhhj.
GN YHHJ OR B3485 OR SF3501 OR S4262.
OS Escherichia coli, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 623;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=94316500; PubMed=8041620;
RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
RT "Analysis of the Escherichia coli genome. V. DNA sequence of the
RT region from 76.0 to 81.5 minutes.";
RL Nucleic Acids Res. 22:2576-2586(1994).
RN [2]
RP SEQUENCE OF 225-374 FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=93259920; PubMed=8387990;
RA Zhao S., Sandt C.H., Peulner G., Vlazny D.A., Gray J.A., Hill C.W.;
RT "Rhs elements of Escherichia coli K-12: complex composites of shared
RT and unique components that have different evolutionary histories.";
RL J. Bacteriol. 175:2799-2808(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=2272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
RN [5]
RP SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
RP (Potential).
CC -! SIMILARITY: Belongs to the ABC-2 integral membrane protein family.
CC -! SIMILARITY: TO E.COLI YBHR AND YBHS AND TO THE C-TERMINAL OF
CC E.COLI YHIG.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP001119; BAB13059.1; -.
CC InterPro; IPR001130; TatD_DNase.
CC Pfam; PF01026; TatD_DNase; 1.
CC TIGRFAMs; TIGR00010; TIGR00010; 1.
CC PROSITE; PS01137; TatD_1; 1.
CC PROSITE; PS01090; TatD_2; 1.
CC PROSITE; PS01091; TatD_3; FALSE_NEG.
KW Hypothetical protein; Hydrolase; Nuclease; Complete proteome.
SQ SEQUENCE 264 AA; 30520 MW; 7f1da900c18e0aac CRC64;
Query Match 29.0%; Score 49; DB 1; Length 264;
Best Local Similarity 45.0%; Pred. No. 8.2;
Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 8 METRTPDINPAWYTGRIIP 27
DB 210 IETDSPYLSFAPYRGKGNQP 229
DR EcoGene; EGI1767; yhhJ
DR InterPro; IPR000412; ABC_transpt2.
DR PROSITE; PS00890; ABC2_MEMBRANE; 1.
KW Hypothetical protein; Transport; Transmembrane; Inner membrane;
Complete proteome.
FT TRANSMEM 23 43 POTENTIAL.
FT TRANSMEM 173 193 POTENTIAL.
FT TRANSMEM 230 250 POTENTIAL.
FT TRANSMEM 256 276 POTENTIAL.
FT TRANSMEM 284 304 POTENTIAL.
FT TRANSMEM 343 363 POTENTIAL.
SQ SEQUENCE 374 AA; 41061 MW; 02895fbl3f493391 CRC64;
Query Match 29.3%; Score 49.5; DB 1; Length 374;
Best Local Similarity 47.4%; Pred. No. 10;
Matches 9; Conservative 5; Mismatches 2; Indels 3; Gaps 1;
QY 7 SMETR---TPDINPAWYTG 22
DB 160 SLETRMRFNFLDPAPWFG 178
RESULT 11
Y355_BUCAI
ID Y355_BUCAI STANDARD; PRT; 264 AA.
AC P57436;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative deoxyribonuclease BU355 (EC 3.1.21.-).
GN BU355.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. APS.";
RL Nature 407:81-86(2000).
CC -! SIMILARITY: BELONGS TO THE TATD DNASE FAMILY. STRONG, TO E.COLI
CC YCFH.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP001119; BAB13059.1; -.
CC InterPro; IPR001130; TatD_DNase.
CC Pfam; PF01026; TatD_DNase; 1.
CC TIGRFAMs; TIGR00010; TIGR00010; 1.
CC PROSITE; PS01137; TatD_1; 1.
CC PROSITE; PS01090; TatD_2; 1.
CC PROSITE; PS01091; TatD_3; FALSE_NEG.
KW Hypothetical protein; Hydrolase; Nuclease; Complete proteome.
SQ SEQUENCE 264 AA; 30520 MW; 7f1da900c18e0aac CRC64;
Query Match 29.0%; Score 49; DB 1; Length 264;
Best Local Similarity 45.0%; Pred. No. 8.2;
Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 8 METRTPDINPAWYTGRIIP 27
DB 210 IETDSPYLSFAPYRGKGNQP 229
```



```

SEQUENCE FROM N.A.
RC STRAIN=XS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RX MEDLINE=22723752; PubMed=12840036;
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens.";
RL Genome Res. 13:1572-1579(2003).
CC
CC -!- FUNCTION: Bidirectionally degrades single-stranded DNA into large
CC acid-insoluble oligonucleotides, which are then degraded further
CC into small acid-soluble oligonucleotides (By similarity).
CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.
CC -!- SUBUNIT: Heterooligomer composed of large and small subunits (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the xseA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF005217; BAC17888.1; ALT_INIT.
CC HAMAP; MF_00378; -; 1.
CC InterPro; IPR003753; Exonuc_VII_L.
CC InterPro; IPR008994; Nucleic_acid_OB.
CC Pfam; PF02601; Exonuc_VII_L; 1.
CC TIGRFAMs; TIGR00237; xseA; 1.
CC Hydrolase; Nuclease; Exonuclease; Complete proteome.
CC KW
CC SEQUENCE 413 AA; 45171 MW; D3BABD3687C6E5D CRC64;
CC
CC Query Match 28.4%; Score 48; DB 1; Length 413;
CC Best Local Similarity 52.4%; Pred. No. 18;
CC Matches 11; Conservative 1; Mismatches 1; Indels 8; Gaps 1;
CC
CC QY 17 PAWYTGKG-----IRPVG 29
CC ||:|||
CC 90 PAFYAGRGSLWVTDIRPVG 110
CC
CC RESULT 15
CC EX7L CORGL STANDARD; PRT; 417 AA.
CC AC QSNW3;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
CC DE (Exonuclease VII large subunit).
CC GN XSEA OR CGLI025.
CC OS Corynebacterium glutamicum (Brevibacterium flavum).
CC OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
CC OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
CC OX NCBI_TaxID=1718;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
CC RA Nakagawa S.;
CC RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
CC Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Bidirectionally degrades single-stranded DNA into large
CC acid-insoluble oligonucleotides, which are then degraded further
CC into small acid-soluble oligonucleotides (By similarity).
CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.
CC -!- SUBUNIT: Heterooligomer composed of large and small subunits (By
CC similarity).

```

```

-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
-!- SIMILARITY: Belongs to the xseA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP005277; BAB98418.1; -.
CC HAMAP; MF_00378; -; 1.
CC InterPro; IPR003753; Exonuc_VII_L.
CC InterPro; IPR008994; Nucleic_acid_OB.
CC Pfam; PF02601; Exonuc_VII_L; 1.
CC TIGRFAMs; TIGR00237; xseA; 1.
CC Hydrolase; Nuclease; Exonuclease; Complete proteome.
CC KW
CC SEQUENCE 417 AA; 45582 MW; B32CD9286C173C34 CRC64;
CC
CC Query Match 28.4%; Score 48; DB 1; Length 417;
CC Best Local Similarity 52.4%; Pred. No. 18;
CC Matches 11; Conservative 1; Mismatches 1; Indels 8; Gaps 1;
CC
CC QY 17 PAWYTGKG-----IRPVG 29
CC ||:|||
CC 94 PAFYAGRGSLWVTDIRPVG 114
CC
CC Search completed: August 12, 2004, 14:44:53
CC Job time : 9.11047 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 42.7151 Seconds
(without alignments)
228.984 Million cell updates/sec

Title: US-09-700-643A-3
Perfect score: 169
Sequence: 1 SRAHQHSEMTPTDINPAWYTGIRPVGRX 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL.25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	157	92.9	98	Q8WN12	Q8wn12 ovis aries
2	101	59.8	117	Q9W624	Q9w624 carassius a
3	68	40.2	692	Q9L8J6	Q9l8j6 rhodospiril
4	64	37.9	1245	Q7X3H6	Q7x3h6 thiocapsa r
5	61	36.1	1236	Q9JPA4	Q9jpa4 rhodocyclus
6	58	34.3	1253	Q8RTV7	Q8rtv7 uncultured
7	58	34.3	1292	Q9ZG55	Q9zg55 heliobacill
8	56	33.1	355	Q95YJ8	Q95yj8 ciona savig
9	54	32.0	314	Q7VSA1	Q7vsa1 bordetella
10	53	31.4	54	Q7Z5Y1	Q7z5y1 homo sapien
11	53	31.4	139	Q8BHP9	Q8bhp9 mus musculu
12	53	31.4	465	Q60587	Q60587 homo sapien
13	53	31.4	465	Q8WW85	Q8ww85 homo sapien
14	53	31.4	625	Q89VA3	Q89va3 bradyrhizob
15	53	31.4	689	Q8X3R1	Q8x3r1 mus musculu
16	53	31.4	767	Q8JZN2	Q8jzn2 mus musculu

Q91W50 mus musculu
Q80TP8 mus musculu
Q8KZ57 uncultured
Q7TtW0 synchococc
Q7TtZ3 prochloroco
Q9hvi8 pseudomonas
Q8tts7 methanosarc
Q8YWC7 anabaena sp
Q8K746 streptococc
Q91SC6 arabidopsis
Q8DM52 synchococc
Q8KRY1 vibrio chol
Q19530 caenorhabdi
Q96SD4 homo sapien
Q8N6C2 homo sapien
Q8IX2 homo sapien
Q8CWS9 mus musculu
Q8SSW8 encephalito
Q890X9 clostridium
Q7VC39 prochloroco
Q8689 saccharomyc
Q84W10 arabidopsis
Q818A7 arabidopsis
Q33480 propionibac
Q8NMC6 corynebacte
Q91FA0 arabidopsis
Q81UV7 homo sapien
Q7UPG7 rhodopirell
Q7WP50 bordetella

ALIGNMENTS

RESULT 1

Q8WN12 PRELIMINARY; PRT; 98 AA.

AC Q8WN12;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Preprolactin-releasing peptide.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Curlewis J.D., Kusters D.H.L., Barclay J.L., Anderson S.T.;
RT "Prolactin-releasing peptide (PrRP) in the ewe: cDNA cloning, mRNA
RT distribution and effects on prolactin secretion in vitro and in
RT vivo."
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF450453; AAL47178.1.;
SQ SEQUENCE 98 AA, 10513 MW, 2A53331E62CAAB5 CRC64;

Query Match 92.9%; Score 157; DB 6; Length 98;
Best Local Similarity 93.3%; Pred. No. 8.7e-16;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRAHQHSEMTPTDINPAWYTGIRPVGR 30

Db 23 SRAHQHSEMTPTDINPAWYTGIRPVGR 52

RESULT 2

Q9W624 PRELIMINARY; PRT; 117 AA.

AC Q9W624;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)

RT "Genes Involved in the Biosynthesis of Photosynthetic Pigments in the
RT Purple Sulfur Photosynthetic Bacterium *Thiocapsa roseopersicina*.";
RT *Ann. Environ. Microbiol.* 69:3093-3102 (2003).

Query Match 37.9%; Score 64; DB 2; Length 1245;

Best Local Similarity 40.6%; Pred. No. 1.5;
Matches 13; Conservative 5; Mismatches 8; Indels 6; Gaps

1101 EQUAT EPTDRTM INIUVFECM KUCVECVBOT 1152

RESULT 5

AC	Q9JPA4;	
DT	01-OCT-2000	(T=EMBL;e) 15 Created)

DI 01-JUN-2003 (IEMBLref. 24, last annotation update)
DE Mg protoporphyrin methyl transferase.

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Comamonadaceae; Rubrivivax.

KN	[1]	RP	SEQUENCE FROM N.A.
1	1	1	1
2	2	2	2
3	3	3	3
4	4	4	4
5	5	5	5
6	6	6	6
7	7	7	7
8	8	8	8
9	9	9	9
10	10	10	10
11	11	11	11
12	12	12	12
13	13	13	13
14	14	14	14
15	15	15	15
16	16	16	16
17	17	17	17
18	18	18	18
19	19	19	19
20	20	20	20
21	21	21	21
22	22	22	22
23	23	23	23
24	24	24	24
25	25	25	25
26	26	26	26
27	27	27	27
28	28	28	28
29	29	29	29
30	30	30	30
31	31	31	31
32	32	32	32
33	33	33	33
34	34	34	34
35	35	35	35
36	36	36	36
37	37	37	37
38	38	38	38
39	39	39	39
40	40	40	40
41	41	41	41
42	42	42	42
43	43	43	43
44	44	44	44
45	45	45	45
46	46	46	46
47	47	47	47
48	48	48	48
49	49	49	49
50	50	50	50
51	51	51	51
52	52	52	52
53	53	53	53
54	54	54	54
55	55	55	55
56	56	56	56
57	57	57	57
58	58	58	58
59	59	59	59
60	60	60	60
61	61	61	61
62	62	62	62
63	63	63	63
64	64	64	64
65	65	65	65
66	66	66	66
67	67	67	67
68	68	68	68
69	69	69	69
70	70	70	70
71	71	71	71
72	72	72	72
73	73	73	73
74	74	74	74
75	75	75	75
76	76	76	76
77	77	77	77
78	78	78	78
79	79	79	79
80	80	80	80
81	81	81	81
82	82	82	82
83	83	83	83
84	84	84	84
85	85	85	85
86	86	86	86
87	87	87	87
88	88	88	88
89	89	89	89
90	90	90	90
91	91	91	91
92	92	92	92
93	93	93	93
94	94	94	94
95	95	95	95
96	96	96	96
97	97	97	97
98	98	98	98
99	99	99	99
100	100	100	100

RT "Phylogenetic analysis of photosynthetic genes of *Rhodocyclus*
RT *gelatinosus*: Possibility of horizontal gene transfer in purple

[2] *Encyclopaedia of Mathematics*, Vol. 10, 1994 (1995).

RX MEDLINE=94132007; PubMed=8300574;
RA Nagashima K.V., Matsuura K., Ohyama S., Shimada K.;

RT	gelatinosus";	260.2477-2484 (1994)
PI	T. Bico Chem	

RP SEQUENCE FROM N.A.
RC STRAIN=IL144;

RT
gelatinous.";
Pl. (In) Garab G. (eds.):

RL mechanisms and effects (Proceedings of the fifth international congress on photosynthesis), pp.4:2889-2892, Kluwer Academic Publishers,

RP	SEQUENCE FROM N.A.
RC	STRAIN=IL14;

Parot P., Vermiglio A.;

RT Rubrivivax gelatinosus. ;
RL Biochemistry 38:15238-15244 (1999).

```
DR PIR; T50904; T50904.
DR GO; GO:00016740; P:transferase activity; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR003672; COB/Mg chel_1.
DR Pfam; PF02514; COB/Mg chel_1.
DR Transferase.
KW Transferase.
SQ SEQUENCE 1236 AA; 134729 MW; 84051C045638520C CRC64;

Query Match 36.1%; Score 61; DB 2; Length 1236;
Best Local Similarity 37.5%; Pred. No. 4.2;
Matches 12; Conservative 6; Mismatches 8; Indels 6; Gaps 1;

QY 3 AHQHSMETRTPDINPAWYTG-----RGIRPV 28
Db 1112 SEQVALETRTRMLNPKWYEGMLKHGYGVGVRQI 1143

RESULT 6
Q8RTV7 PRELIMINARY; PRT; 1253 AA.
AC Q8RTV7;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE COB/magnesium chelate family protein.
GN BBAC000-65D09.24.
OS uncultured proteobacterium.
OC Bacteria; Proteobacteria; environmental samples.
OX NCBI_TaxID=153809;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21822632; PubMed=11832943;
RA Beja O., Suzuki M.T., Heideberg J.F., Nelson W.C., Preston C.M.,
RA Hamada T., Eisen J.A., Fraser C.M., DeLong E.F.;
RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL Nature 415:630-633(2002).
DR EMBL; AF008919; AAL76369.1; -.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR003672; COB/Mg chelase.
DR Pfam; PF02514; COB/Mg chel_1.
SQ SEQUENCE 1253 AA; 137586 MW; C98ABC010E261511 CRC64;

Query Match 34.3%; Score 58; DB 2; Length 1253;
Best Local Similarity 37.5%; Pred. No. 12;
Matches 12; Conservative 5; Mismatches 9; Indels 6; Gaps 1;

QY 3 AHQHSMETRTPDINPAWYTG-----RGIRPV 28
Db 1129 AEQVALETRTRVLPNPKWYESMLDHYGVGVRAI 1160

RESULT 7
Q9ZGES PRELIMINARY; PRT; 1292 AA.
AC Q9ZGES;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE MG chelate subunit H Bchl.
GN Bchl.
OS Helicobacillus mobilis.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Helicobacteriaceae;
OX Helicobacillus.
OX NCBI_TaxID=29064;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99061957; PubMed=98439979;
RA Xiong J., Inoue K., Bauer C.E.;
RT "Tracking molecular evolution of photosynthesis by characterization of
a major photosynthesis gene cluster from Helicobacillus mobilis.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:14851-14856(1998).
DR EMBL; AF080002; AAC84033.1; -.
DR PIR; T31462; T31462.
```

RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 RT Bordetella parapertussis and Bordetella bronchiseptica.",
 RL Nat. Genet. 35:32-40 (2003).
 DR EMBL; BX640412; CAB44866.1; -.
 KW Hydrolase; Complete proteome.
 SQ SEQUENCE 314 AA; 33764 MW; A90870683CSBE0B4 CRC64;

Query Match 32.0%; Score 54; DB 16; Length 314;
 Best Local Similarity 54.5%; Pred. No. 9.9;
 Matches 12; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 7 SMETRTPDINPAWYTGIRPV 28
 Db 164 SMAGRTPAIEPGWVRQIGIRSV 185

RESULT 10

Q7Z6Y1 PRELIMINARY; PRT; 54 AA.

ID Q7Z6Y1
 AC Q7Z6Y1
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE D479J7.3 (Sushi-repeat protein (SRPUL)) (Fragment).
 GN D479J7.3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Lawlor S.;
 RP Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AL035608; CAB55682.1; -.
 DR NON_TER 54
 FT SEQUENCE 54 AA; 6110 MW; E2F3C39F7B961A9F CRC64;

Query Match 31.4%; Score 53; DB 4; Length 54;
 Best Local Similarity 50.0%; Pred. No. 1.9;
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 12 TPDINPAWYTGIRPV 27
 Db 18 TPAVPTWYAGSGYYP 33

RESULT 11

Q8BHP9 PRELIMINARY; PRT; 139 AA.

ID Q8BHP9
 AC Q8BHP9
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Weakly similar to hypothetical protein KIAA0574.
 GN 5730507A09RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Body;
 RX MEDLINE=22354683; PubMed=12466851;
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.",
 RL Nature 420:563-573 (2002).
 DR EMBL; AK017751; BAC25529.1; -.

DR MGD; MGI:1917888; 5730507A09RIK.
 KW Hypothetical protein.
 SQ SEQUENCE 139 AA; 14740 MW; A08DD1B09441B259 CRC64;

Query Match 31.4%; Score 53; DB 11; Length 139;
 Best Local Similarity 50.0%; Pred. No. 5.6;
 Matches 12; Conservative 2; Mismatches 6; Indels 4; Gaps 2;

Qy 4 HQHSEMTPTDINPAWYTGIRPV 27
 Db 47 HQHS-STGDPD---TWKTGQTKP 66

RESULT 12

Q06087 PRELIMINARY; PRT; 465 AA.

ID Q06087
 AC Q06087
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Sushi-repeat protein (Sushi-repeat containing protein).
 GN SRPUL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kurosawa H., Inukai T., Inaba T., Goi K., Chang K.-S., Sinjo T.,
 RA Rakeshaw K.M., Naeve C.W., Lock T.A.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.

RA Huang C.-H., Chen H., Peng J., Chen Y.;
 RT "Cloning and characterization of the sushi-repeat containing protein
 RT (SRP) as a novel interaction partner of Rn type C glycoprotein
 RT (RHC)."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF060567; AAC15765.1; -.
 DR EMBL; AF393649; AAM73693.1; -.
 GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR InterPro; IPR003410; Hyalin.
 DR InterPro; IPR000436; Sushi_SCR_CCP.
 DR Pfam; PF02494; HYR; 1.
 DR Pfam; PF00084; sushi; 3.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 SQ SEQUENCE 465 AA; 52971 MW; 4D752B187FF3EFB8 CRC64;

Query Match 31.4%; Score 53; DB 4; Length 465;
 Best Local Similarity 50.0%; Pred. No. 22;
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 12 TPDINPAWYTGIRPV 27
 Db 18 TPAVPTWYAGSGYYP 33

RESULT 13

Q8W85 PRELIMINARY; PRT; 465 AA.

ID Q8W85
 AC Q8W85
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Sushi-repeat protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -|- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
DR EMBL; BC024826; AAH24826.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004601; F:peroxidase activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006979; P:response to oxidative stress; IEA.
DR InterPro; IPR002059; Cold shock.
DR InterPro; IPR008994; Nucleic acid_OB.
DR InterPro; IPR002016; Peroxidase.
DR Pfam; PF00313; CSD; 6.
DR ProDom; PD000621; Cold shock; 1.
DR SMART; SM00357; CSP_4.
DR PROSITE; PS00352; COLD_SHOCK; 3.
DR PROSITE; PS00436; PEROXIDASE_2; 1.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 689 AA; 76861 MW; 470743A80947A3A9 CRC64;
Query Match 31.4%; Score 53; DB 11; Length 689;
Best Local Similarity 43.5%; Pred. No. 34;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;
QY 6 HSMETRTPDINPAWYTGIRPV 28
DB 474 HSVGITEEANTPTIYSGKVRPL 496
Search completed: August 12, 2004, 14:49:03
Job time : 43.7151 secs

RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020733; AAH20733.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR003410; Hvalin.
DR InterPro; IPR000436; Sush1_SCR_CCP.
DR Pfam; PF02494; HVR; 1.
DR Pfam; PF00084; sush1; 3.
DR SMART; SM00032; CCP; 3.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
SQ SEQUENCE 465 AA; 52957 MW; 3D7229487DA1B8BD CRC64;
Query Match 31.4%; Score 53; DB 4; Length 465;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 12 TPDINPAWYTGIRPV 27
DB 18 TPVPTWVAGSGYVP 33
RESULT 14
Q89VA3 PRELIMINARY; PRT; 625 AA.
AC Q89VA3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Bll1144 protein.
GN Bll1144.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=2248498; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Iidesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shampo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AF005939; BAC46409.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003910; F:DNA ligase (ATP) activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR000977; DNA_ligase.
DR Pfam; PF01068; DNA_ligase; 1.
DR Pfam; PF04679; DNA_ligase_A; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; 1.
DR PROSITE; PS0160; DNA_LIGASE_A3; 1.
KW Complete proteome.
SQ SEQUENCE 625 AA; 69415 MW; 56BED3A80A46AD42 CRC64;
Query Match 31.4%; Score 53; DB 16; Length 625;
Best Local Similarity 38.7%; Pred. No. 31;
Matches 12; Conservative 3; Mismatches 10; Indels 6; Gaps 1;
QY 6 HSMETRTPDINP-----AWYTGIRPVGR 30
DB 240 HEVELIWPGLAPFYLDLFAWLEGRGKPVNR 270
RESULT 15
Q8R3R1 PRELIMINARY; PRT; 689 AA.
ID Q8R3R1
AC Q8R3R1;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 65.7849 Seconds
(without alignments)
133.146 Million cell updates/sec

Title: US-09-700-643A-3
Perfect score: 169
Sequence: 1 SRAHQSMETRPDINPAWYTGIRPVGRX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	168	99.4	31	2 AAW31384	AAW31384 Rat type
2	168	99.4	31	2 AAW95174	AAW95174 Murine pi
3	168	99.4	31	2 AAW95173	AAW95173 Murine pi
4	168	99.4	31	2 AAW87614	AAW87614 Rat 19p2
5	168	99.4	31	2 AAW97233	AAW97233 Rat type
6	168	99.4	31	3 AAY49292	AAY49292 19p2 liga
7	168	99.4	31	3 AAY87504	AAY87504 Rat prola
8	168	99.4	31	3 AAB10355	AAB10355 Rat oxyto
9	168	99.4	31	4 AAB90993	AAB90993 Prolactin
10	168	99.4	31	4 AAG62524	AAG62524 Rat CRH r
11	168	99.4	31	5 AAE28400	AAE28400 Rat PRP
12	168	99.4	31	6 AAEU60826	AAEU60826 Peptide p
13	168	99.4	31	6 ABUE60837	ABUE60837 Peptide p
14	168	99.4	32	2 AAW31385	AAW31385 Rat oxyto
15	168	99.4	32	3 AAB10356	AAB10356 Rat oxyto
16	168	99.4	32	4 AAG62525	AAG62525 Rat CRH r
17	168	99.4	32	6 ABUE60838	ABUE60838 Peptide p
18	168	99.4	33	2 AAW31386	AAW31386 Rat type
19	168	99.4	33	3 AAB10357	AAB10357 Rat oxyto
20	168	99.4	33	4 AAG62526	AAG62526 Rat CRH r
21	168	99.4	33	6 ABUE60839	ABUE60839 Peptide p
22	168	99.4	82	2 AAW95172	AAW95172 Murine pi
23	168	99.4	83	2 AAW31383	AAW31383 Rat type
24	168	99.4	83	2 AAW97225	AAW97225 Rat type
25	168	99.4	83	3 AAB10354	AAB10354 Rat oxyto

26	168	99.4	83	4 AAG62523	AAG62523 Rat CRH r
27	158	93.5	31	4 AAW73370	AAW73370 bFRRp31 P
28	157	92.9	31	2 AAW31371	AAW31371 Bovine G
29	157	92.9	31	2 AAW95188	AAW95188 Bovine pi
30	157	92.9	31	2 AAW87613	AAW87613 Bovine 19
31	157	92.9	31	2 AAW97218	AAW97218 Bovine pi
32	157	92.9	31	3 AAY49298	AAY49298 19p2 liga
33	157	92.9	31	3 AAY49290	AAY49290 19p2 liga
34	157	92.9	31	4 AAB10347	AAB10347 Bovine CR
35	157	92.9	31	4 AAG62516	AAG62516 Bovine CR
36	157	92.9	31	5 AAE26399	AAE26399 Bovine Pr
37	157	92.9	31	6 ABUE60825	ABUE60825 Peptide p
38	157	92.9	31	6 ABUE60831	ABUE60831 Peptide p
39	157	92.9	32	2 AAW31372	AAW31372 Bovine G
40	157	92.9	32	2 AAW95189	AAW95189 Bovine pi
41	157	92.9	32	3 AAB10348	AAB10348 Bovine CR
42	157	92.9	32	4 AAG62517	AAG62517 Bovine CR
43	157	92.9	32	6 ABUE60832	ABUE60832 Peptide p
44	157	92.9	33	2 AAW31373	AAW31373 Bovine G
45	157	92.9	33	2 AAW95190	AAW95190 Bovine pi

ALIGNMENTS

RESULT 1

AAW31384
ID AAW31384 standard; peptide; 31 AA.
XX
AC AAW31384;
XX
DT 27-AUG-2003 (revised)
DT 06-APR-1998 (first entry)
XX
DE Rat type G protein-coupled receptor ligand fragment 1.
XX
KW G protein-coupled receptor; ligand binding; pharmacological; modulator;
KW pituitary; central nervous system; pancreas; prophylactic;
KW therapeutic agent.
XX
OS Rattus sp.
XX
PN WO9724436-A2.
XX
PD 10-JUL-1997.
XX
PF 26-DEC-1996; 96WO-JP003821.
XX
PR 28-DEC-1995; 95JP-00343371.
PR 15-MAR-1996; 96JP-00059419.
PR 12-AUG-1996; 96JP-00211805.
PR 18-SEP-1996; 96JP-00246573.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;
PI Kitada C;
XX
WPI: 1997-363672/33.
DR N-PSDB; AAV02421.
XX
PT Ligand peptide for G protein-coupled receptor - acts by modulating
PT function in the central nervous system, pancreas and pituitary gland.
PS
PS Claim 2; Page 179; 258pp; English.
XX
CC This sequence represents a peptide fragment from a novel rat type ligand
CC polypeptide corresponding to amino acid residues 22 to 52 of the sequence
CC represented in AAW31383 and is used in an assay to monitor ligand binding
CC to the G protein-coupled receptor protein. Pharmacological compositions
CC containing this ligand may be used as a pituitary function modulator, a
CC central nervous system modulator or a pancreatic function modulator. This
CC ligand could have specific applications as a prophylactic or therapeutic

CC agent for dementia, depression, hyperkinetic syndrome, disturbance of
 CC consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone
 CC secretory disease, hyper- and polyphagia, hyperlipidaemia,
 CC hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,
 CC amyotrophic lateral sclerosis, acute myocardial infarction, infertility,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis and/or oligosaccharitis. Assays can also be developed to screen
 CC compounds which are capable of altering the binding activity of the
 CC ligand affecting activation of the G protein-coupled receptor protein.
 CC (Updated on 27-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 31 AA;

Query Match 99.4%; Score 168; DB 2; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.5e-17;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRAHQSMETPTDINPAWYTGIRPVGR 30
 |||||
 DB 1 SRAHQSMETPTDINPAWYTGIRPVGR 30
 |||||

RESULT 2
 AAW95174
 ID AAW95174 standard; protein; 31 AA.
 XX
 AC AAW95174;
 XX
 DT 10-MAR-1999 (first entry)
 XX
 DE Murine pituitary-derived ligand polypeptide antigenic epitope.
 XX
 KW Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
 KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
 KW Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
 KW gene therapy; transgenic animal; epitope.
 XX
 OS Mus sp.
 XX
 FN WO9849295-A1.
 XX
 PD 05-NOV-1998.
 XX
 PF 27-APR-1998; 98WO-JP001923.
 XX
 PR 28-APR-1997; 97JP-00109974.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Hinuma S, Fukusumi S;
 XX
 DR WPI; 1999-009423/01.
 XX

PT New polypeptide ligand for orphan G protein coupled receptors - used for
 PT treating disorders of central nervous system, pituitary and pancreas, and
 PT for drug screening.
 PS Disclosure; Page 26; 206pp; English.

XX The invention relates to a murine pituitary-derived ligand polypeptide
 CC which is a ligand for the G-protein coupled orphan receptor designated
 CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing
 CC the ligand polypeptide encoding DNA are used to produce a recombinant
 CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate
 CC function of the pituitary, central nervous system, pancreas and other
 CC tissues and can be used to screen for agents that modulate binding of the
 CC polypeptide to the receptor; to quantify the amount of receptor in a
 CC sample and to raise antibodies. They may also be used therapeutically.

CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's
 CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;
 CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;
 CC rheumatoid arthritis, epilepsy and many others, also to improve post-
 CC operative nutritional status and as vasopressor. Transgenic animals
 CC carrying the ligand polypeptide encoding DNA or its mutant are used to
 CC study the function of the polypeptide-expressing genes, as models of
 CC disease, for drug screening and as source of cell lines. The ligand
 CC polypeptide DNA is used as a source of probes and primers; to identify
 CC related sequences; in receptor-binding assays; for production of Ab and
 CC antisera; in drug development; for gene therapy and to develop transgenic
 CC animals. Sequences AAW95174 to AAW95178 represent antigenic epitopes
 CC which can be used for the preparation of anti-ligand polypeptide antibody
 XX
 SQ Sequence 31 AA;

Query Match 99.4%; Score 168; DB 2; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.5e-17;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRAHQSMETPTDINPAWYTGIRPVGR 30
 |||||
 DB 1 SRAHQSMETPTDINPAWYTGIRPVGR 30
 |||||

RESULT 3
 AAW95173
 ID AAW95173 standard; peptide; 31 AA.
 XX
 AC AAW95173;
 XX
 DT 10-MAR-1999 (first entry)
 XX
 DE Murine pituitary-derived ligand mature polypeptide sequence.
 XX
 KW Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
 KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
 KW Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
 KW gene therapy; transgenic animal.
 XX
 OS Mus sp.
 XX
 FN WO9849295-A1.
 XX
 PD 05-NOV-1998.
 XX
 PF 27-APR-1998; 98WO-JP001923.
 XX
 PR 28-APR-1997; 97JP-00109974.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Hinuma S, Fukusumi S;
 XX
 DR WPI; 1999-009423/01.
 XX

PT New polypeptide ligand for orphan G protein coupled receptors - used for
 PT treating disorders of central nervous system, pituitary and pancreas, and
 PT for drug screening.
 PS Disclosure; Page 134; 206pp; English.

XX This represents the matured murine pituitary-derived ligand polypeptide
 CC sequence. The polypeptide is a ligand for the G-protein coupled orphan
 CC receptor designated GPR10 (human) or UHR-1 (rat). Cells transformed with
 CC a vector containing the ligand polypeptide encoding DNA are used to
 CC produce a recombinant ligand polypeptide. The ligand polypeptide, and its
 CC fragments, modulate function of the pituitary, central nervous system,
 CC pancreas and other tissues and can be used to screen for agents that
 CC modulate binding of the polypeptide to the receptor; to quantify the

CC amount of receptor in a sample and to raise antibodies. They may also be
 CC used therapeutically, e.g. to treat senile dementia; Alzheimer's,
 CC Parkinson's or Huntington's diseases; Creutzfeldt-Jakob disease; poisoning
 CC by heavy metals or drugs; diabetes; schizophrenia; disorders of growth
 CC hormone secretion; cancer; rheumatoid arthritis, epilepsy and many
 CC others, also to improve post-operative nutritional status and as
 CC vasopressor. Transgenic animals carrying the ligand polypeptide encoding
 CC DNA or its mutain are used to study the function of the polypeptide-
 CC expressing genes, as models of disease, for drug screening and as source
 CC of cell lines. The ligand polypeptide DNA is used as a source of probes
 CC and primers; to identify related sequences; in receptor-binding assays;
 CC for production of Ab and antisera; in drug development; for gene therapy
 CC and to develop transgenic animals
 CC
 XX
 SQ Sequence 31 AA;

Query Match 99.4%; Score 168; DB 2; Length 31;
 Best Local Similarity 100.0%; Pred. NO. 1.5e-17;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQHSMETRPDINPAWYTGIRPVGR 30
 |||||
 Db 1 SRAHQHSMETRPDINPAWYTGIRPVGR 30

RESULT 4
 AAW87614
 ID AAW87614 standard; peptide; 31 AA.

XX
 AC AAW87614;

DT 29-MAR-1999 (first entry)

DE Rat 19P2 ligand.

KW 19P2 ligand; G protein coupled receptor; pituitary;
 KW prolactin releasing peptide; rat; dementia; breast cancer; therapy.

XX Rattus sp.

PN EP887417-A2.

PD 30-DEC-1998.

PF 25-JUN-1998; 98EP-00111725.

PR 27-JUN-1997; 97JP-00172118.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Suenaga M, Moriya T, Tanaka Y, Nishimura O;

DR WPI; 1999-047884/05.

PT Producing a 19P2 pituitary G protein receptor ligand - by cleavage of a
 PT fusion protein, useful for preventing and treating dementia, breast
 PT cancer, renal failure and autoimmune disease.

PS Claim 5; Page 34; 56pp; English.

CC This is the amino acid sequence of the rat pituitary G protein-coupled
 CC receptor ligand 19P2L. A method suitable for commercial high-level
 CC production of 19P2L comprises expressing the ligand in host cells as a
 CC recombinant fusion protein e.g. with human basic fibroblast growth factor
 CC (see AAV83794-95) that has been modified to include an N-terminal
 CC cysteine residue. The ligand is released from the fusion by cyanylation
 CC followed by ammonolysis. 19P2L has prolactin secretion-stimulating and
 CC (at high doses) prolactin secretion-inhibiting properties. It can be used
 CC in the treatment and prevention of various diseases including: senile
 CC dementia, cerebrovascular dementia, and dementia associated with:
 CC neurological disorders (e.g. Alzheimer's disease, Parkinson's disease,
 CC Pick's disease, Huntington's disease), infectious diseases (e.g.
 CC Creutzfeldt-Jakob's), endocrine or metabolic disease or toxicosis (e.g.

CC hypothyroidism, vitamin B12 deficiency, alcoholism, intoxication by
 CC drugs, metal and organic compounds), tumorigenic diseases (e.g. brain
 CC tumour), traumatic diseases (e.g. chronic subarachnoid haemorrhage, and
 CC other types of dementia, depression, hyperactive child syndrome
 CC (microencephalopathy) and disturbance of consciousness. It is also useful
 CC for prevention and treatment of diseases associated with prolactin hypo
 CC and hypersecretion respectively, including: hyperprolactinaemia,
 CC pituitary adenoma, breast cancer, infertility, impotence and autoimmune
 CC disease (hypersecretion disorders), and seminal vesicle hypoplasia,
 CC osteoporosis, menopausal syndrome and renal failure (hypersecretion
 CC disorders). The 19P2 polypeptide/amide is also useful as a test reagent
 CC for study of the prolactin secretory function or as a lactagogue in
 CC mammalian farm animals
 CC
 XX
 SQ Sequence 31 AA;

Query Match 99.4%; Score 168; DB 2; Length 31;
 Best Local Similarity 100.0%; Pred. NO. 1.5e-17;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQHSMETRPDINPAWYTGIRPVGR 30
 |||||
 Db 1 SRAHQHSMETRPDINPAWYTGIRPVGR 30

RESULT 5
 AAW97233
 ID AAW97233 standard; peptide; 31 AA.

XX
 AC AAW97233;

DT 06-MAY-1999 (first entry)

DE Rat type ligand polypeptide fragment.

KW Rat type ligand; modulation; prolactin secretion;
 KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;
 KW menopausal syndrome; euthyroid; hypometabolism; lactation;
 KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;
 KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;
 KW acromegaly; Chiari-Prommel syndrome; Argonz-del Castillo syndrome;
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;
 KW contraceptive; placental function; choriocarcinoma; hydatid mole;
 KW interruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;
 KW abnormal lipidmetabolism; oxytocia.

XX Rattus sp.

PN WO9858962-A1.

PD 30-DEC-1998.

PF 22-JUN-1998; 98WO-JP002765.

PR 23-JUN-1997; 97JP-00165437.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Hinuma S, Kawamata Y, Fujii R, Matsumoto H;

DR WPI; 1999-105614/09.

PT Use of G protein-coupled receptor ligands - for modulating prolactin
 PT secretion or placental function, e.g. for treating menopausal syndrome,
 PT tumours, autoimmune disease or abnormal pregnancy.

PS Claim 3; Page 153; 241pp; English.

CC The present sequence represents a rat type ligand fragment. It is used in
 CC the course of the invention. The specification describes an agent for
 CC modulating prolactin secretion which comprises a ligand polypeptide or a
 CC salt, for a G protein-coupled receptor (GPCR) protein. The agents for
 CC promoting prolactin secretion can be used for treating or preventing

CC hypocoovarianism, gonocyst cacogenesis, menopausal syndrome, euthyroid or
 CC hypometabolism. They can be used for promoting lactation in a domestic
 CC mammal and as an aphrodisiac. The agents for inhibiting prolactin
 CC secretion can be used for treating or preventing pituitary adenomatosis,
 CC brain tumour, emmenopathy, autoimmune disease, prolactinoma,
 CC infertility, impotence, amenorrhea, galactorrhea, acromegaly, Chiari-
 CC Frommel syndrome, Argonz-del Castillo syndrome, Forbes-Albright syndrome,
 CC lymphoma, Sheehan syndrome or dyszoospermia. The inhibitory agents can
 CC also be used as contraceptives. The agents for modulating placental
 CC function can be used for treating or preventing choriocarcinoma, hydatid
 CC mole, irruption mole, abortion, unthrifty fetus, abnormal
 CC saccharometabolism, abnormal lipidmetabolism or oxytocia
 XX
 SQ Sequence 31 AA;

Query Match 99.4%; Score 168; DB 2; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.5e-17; Gaps 0;
 Matches 30; Conservative 0; Mismatches 0; Indels 0;

Qy 1 SRAHQHSMETRTDINPAWYTGIRPVGR 30
 |||||
 Db 1 SRAHQHSMETRTDINPAWYTGIRPVGR 30
 |||||

RESULT 6

AA49292
 ID AAY49292 standard; peptide; 31 AA.

AC AAY49292;

DT 22-FEB-2000 (first entry)

XX 19P2 ligand peptide fragment.

XX Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;
 KW pituitary; regulatory mechanism; central nervous system; pancreatic.
 XX

OS Rattus sp.

FH Key Location/Qualifiers
 FT Modified-site 31
 FT /note= "C-terminal amide"

XX WO9960112-A1.

XX 25-NOV-1999.

XX 20-MAY-1999; 99WO-UP002650.

XX 21-MAY-1998; 98JP-00140293.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Matsumoto H, Kitada C, Hinuma S;

XX WPI; 2000-039381/03.

XX New monoclonal antibodies, useful in diagnosis, as drugs and in studying
 PT diseases related to ligand abnormality.

XX Disclosure; Page 26; 73pp; Japanese.

XX The invention provides a monoclonal antibody which has a specific
 CC reaction with the part peptide of the C-terminal of 19P2 ligand or its
 CC derivative. The antibodies can be used in diagnosis or to treat or
 CC prevent diseases associated with abnormality in the pituitary function
 CC regulatory mechanism (e.g. promotion of prolactin secretion), central
 CC nervous regulatory mechanism, and pancreatic function regulatory
 CC mechanism. The antibody-based immunoassay can also be applied in
 CC clarifying the physiological functions of the ligand and its derivative.
 CC Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand

XX Sequence 31 AA;

Query Match 99.4%; Score 168; DB 3; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.5e-17; Gaps 0;
 Matches 30; Conservative 0; Mismatches 0; Indels 0;

Qy 1 SRAHQHSMETRTDINPAWYTGIRPVGR 30
 |||||
 Db 1 SRAHQHSMETRTDINPAWYTGIRPVGR 30
 |||||

RESULT 7

AA87504
 ID AAY87504 standard; protein; 31 AA.

AC AAY87504;

DT 18-JUL-2000 (first entry)

XX Rat prolactin-releasing peptide, PrRP.

XX Prolactin-releasing peptide; PrRP; GPR10; G protein-coupled receptor;
 KW feeding behaviour; food intake; modulation; antagonist; anorectic;
 XX obesity; agonist; cachexia.

OS Rattus sp.

FH Key Location/Qualifiers
 FT Modified-site 31
 FT /note= "C-terminal amide"

XX WO200017641-A1.

XX 30-MAR-2000.

XX 22-SEP-1999; 99WO-US021243.

XX 22-SEP-1998; 98US-0101380P.

XX 14-OCT-1998; 98US-00172353.

XX (MILL-) MILLENNIUM PHARM INC.

XX Stricker-Kongrad A, Gu W;

XX WPI; 2000-303231/26.

XX Identifying modulators of body weight by a combination of a cell-free or
 FT cell-based assay to identify modulators of GPR10, followed by an in vivo
 FT assay for the compounds effect on e.g. feeding behavior.

XX Example 2; Page 61; 82pp; English.

XX The invention relates to a method for identifying compounds useful for
 CC modulating body weight. The method comprises cell-free and/or cell-based
 CC assays that identify compounds which bind to and/or activate or inhibit
 CC the activity of GPR10, a G protein-coupled receptor. These assays are
 CC then followed by an in vivo assay of the effect of the compound on
 CC feeding behaviour, body weight or metabolic rate in a mammal. Prolactin-
 CC releasing peptide (PrRP; AAY87504) is a ligand of GPR10. Binding of PrRP
 CC to GPR10 stimulates a signal transduction cascade, which results in an
 CC increase in food intake. Compounds identified using the method of the
 CC invention are useful for the modulation of body weight. Antagonists of
 CC GPR10 can be used to treat obesity, while GPR10 agonists can be used to
 CC treat cachexia. The present sequence represents rat PrRP

XX Sequence 31 AA;

Query Match 99.4%; Score 168; DB 3; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.5e-17;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQHSMETRTDINPAWYTGIRPVGR 30
 |||||
 Db 1 SRAHQHSMETRTDINPAWYTGIRPVGR 30
 |||||

RESULT 8
 AAB10355
 ID AAB10355 standard; peptide; 31 AA.
 XX AC AAB10355;
 XX DT 24-NOV-2000 (first entry)
 XX DE Rat oxytocin secretion promoting peptide SEQ ID NO: 18.
 XX XX Rat; oxytocin secretion promoter; G protein-coupled receptor protein;
 XX XX treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 XX XX caesarean section; artificial fertilization; galactostasis; goat; pig;
 XX XX veterinary medicine; milk production.
 XX OS Rattus sp.
 XX FN WO200038704-A1.
 XX PD 06-JUL-2000.
 XX PF 22-DEC-1999; 99WO-JP007199.
 XX PR 25-DEC-1998; 98JP-00369585.
 XX XX (TAXE) TAKEDA CHEM IND LTD.
 XX PA Matsumoto H, Kitada C, Hinuma S;
 XX PI WPI; 2000-452298/39.
 XX PT Physiologically-active polypeptide recognized as ligand by G protein-
 PT coupled receptor protein, for promoting secretion of oxytocin, as drugs
 PT for diseases relating to oxytocin secretion and in veterinary medicine.
 XX PS Claim 3; Page 57; 72pp; Japanese.
 XX CC This invention describes a novel oxytocin secretion-regulating agent
 CC which contains a ligand peptide or its salt for the G protein-coupled
 CC receptor protein. It is useful in the form of drugs for ameliorating,
 CC preventing and treating diseases relating to oxytocin secretion e.g. weak
 CC pains and atonic bleeding, before and after expulsion of placenta,
 CC uterine recovery failure, caesarean section, stoppage of artificial
 CC fertilization or galactostasis and is also applicable in veterinary
 CC medicine for promoting milk production in cow, goat and pig. This
 CC sequence represents a rat peptide which acts as an oxytocin secretion
 CC promoter
 XX SQ Sequence 31 AA;
 Query Match 99.4%; Score 168; DB 3; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.5e-17;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRAHQSMETRTDPINPAWYTGIRPVGR 30
 Db 1 SRAHQSMETRTDPINPAWYTGIRPVGR 30
 RESULT 9
 ID AAB90993
 XX DT 22-JUN-2001 (first entry)
 XX DE Prolactin releasing peptide SEQ ID NO:167.
 XX XX Protection; endogenous therapeutic peptide; peptidase; conjugation;
 XX FN blood component; modification; succinimidyl; maleimido group; amino;
 KW

KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
 XX OS Homo sapiens.
 OS Synthetic.
 XX FN WO200069900-A2.
 XX PD 23-NOV-2000.
 XX PF 17-MAY-2000; 2000WO-US013576.
 XX PR 17-MAY-1999; 99US-0134406P.
 PR 10-SEP-1999; 99US-0153406P.
 PR 15-OCT-1999; 99US-0159783P.
 XX PA (CONJ-) CONJUCHEM INC.
 XX PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
 XX WPI; 2001-112059/12.
 XX PT Modifying and attaching therapeutic peptides to albumin prevents
 PT peptidase degradation, useful for increasing length of in vivo activity.
 XX PS Disclosure; Page 244; 733pp; English.
 XX CC The present invention describes a modified therapeutic peptide (I)
 CC comprising a therapeutically active amino acid region (III) and a
 CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
 CC a less therapeutically active amino acid region (IV), which covalently
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
 CC factors and neurotransmitters, to protect them from peptidase activity in
 CC vivo for the treatment of various disorders. Endogenous therapeutic
 CC peptides are not suitable as drug candidates as they require frequent
 CC administration due to rapid degradation by peptidases in the body.
 CC Modifying and attaching therapeutic peptides to albumin prevents or
 CC reduces the action of peptidases to increase length of activity (half
 CC life) and specificity as bonding to large molecules decreases
 CC intracellular uptake and interference with physiological processes.
 CC AAB90829 to AAB92441 represent peptides which can be used in the
 CC exemplification of the present invention
 XX SQ Sequence 31 AA;
 Query Match 99.4%; Score 168; DB 4; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.5e-17;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRAHQSMETRTDPINPAWYTGIRPVGR 30
 Db 1 SRAHQSMETRTDPINPAWYTGIRPVGR 30
 RESULT 10
 ID AAG62524 standard; peptide; 31 AA.
 XX AC AAG62524;
 XX DT 24-AUG-2001 (first entry)
 XX DE Rat CRH releasing protein related peptide SEQ ID NO: 18.
 XX XX Rat; corticotrophin releasing hormone; CRH; G protein receptor ligand;
 KW analgesic; hyperaldosteronism; hypercortislaemia; hypoadrenocorticism;
 KW Addison's disease; adrenal gland hyperfunction; obesity.
 XX OS Rattus sp.
 XX FN WO200035984-A1.

PD 25-MAY-2001.
 XX 17-NOV-2000; 2000WO-JP008119.
 XX 18-NOV-1999; 99JP-00327900.
 PR 26-SEP-2000; 2000JP-00297073.
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX Kitada C, Matsumoto H, Hinuma S;
 XX WPI; 2001-355552/37.
 DR Use of G protein receptor ligand or peptide for controlling corticotropin
 XX releasing hormone secretion.
 XX Claim 3; Page 69; 90pp; Japanese.
 PS The present sequence describes a method of controlling the secretion of
 CC corticotropin releasing hormone (CRH), involving the use of a G protein
 CC receptor ligand. This can be used to control the secretion of CRH and is
 CC useful as an analgesic or for treating, preventing or ameliorating
 CC diseases associated with CRH secretion such as hyperaldosteronism,
 CC hypercortislaemia, secondary or chronic hypoadrenocorticism, Addison's
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The
 CC present sequence is a peptide used in the exemplification of the
 CC invention
 XX Sequence 31 AA;
 SQ
 Query Match 99.4%; Score 168; DB 4; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.5e-17;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRAHQSMETRTDINPAWYTGIRPVGR 30
 DB 1 SRAHQSMETRTDINPAWYTGIRPVGR 30
 RESULT 11
 AAEE26400
 ID AAEE26400 standard; peptide; 31 AA.
 AC AAEE26400;
 XX 13-DEC-2002 (first entry)
 XX Rat PrRP-31 peptide.
 XX Rat; wakefulness; sleep disorder; prolactin releasing peptide receptor;
 XX PrRP; GPR10; therapy; epilepsy; narcolepsy; sleep apnoea;
 XX insomnia; idiopathic hypersomnia; psychogenic hypersomnia; seizure;
 XX anticonvulsant.
 XX Rattus sp.
 XX US2002037533-A1.
 XX 28-MAR-2002.
 XX 17-AUG-2001; 2001US-00932161.
 XX 28-APR-2000; 2000US-00560915.
 XX (CIVE/) CIVELLI O.
 XX (LINS/) LIN S.
 XX Civelli O, Lin S;
 XX WPI; 2002-403931/43.
 XX Screening for compounds useful for promoting wakefulness or sleep, and
 PT for treating sleeping disorders, e.g. insomnia, hypersomnia or sleep
 PT apnea, comprises administering a prolactin releasing peptide agonist or
 PT antagonist.
 XX Disclosure; Page 24; 35pp; English.
 PS The present invention relates to a method of screening for compounds for
 CC promoting wakefulness or sleep in a mammal. The method involves
 CC administering a prolactin releasing peptide (PrRP) receptor (GPR10)
 CC agonist or antagonist respectively and determining the ability of the
 CC compound to promote wakefulness or sleep. The compounds identified from
 CC the method are used in the therapy of epilepsy and other diseases
 CC associated with absence seizures and in promoting wakefulness and sleep
 CC in individuals having sleep disorders such as insomnia and narcolepsy.
 CC PrRP receptor agonists may be used to treat common disorders which lead
 CC to sleepiness, e.g. sleep apnoea, narcolepsy, idiopathic hypersomnia and
 CC psychogenic hypersomnia. PrRP receptor antagonists are useful for
 CC promoting sleep and for treating insomnia such as adjustment sleep
 CC disorder and psychophysiological insomnia. The present sequence is rat PrRP
 CC -31 peptide
 XX Sequence 31 AA;
 SQ
 Query Match 99.4%; Score 168; DB 5; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.5e-17;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRAHQSMETRTDINPAWYTGIRPVGR 30
 DB 1 SRAHQSMETRTDINPAWYTGIRPVGR 30
 RESULT 12
 ABU60826
 ID ABU60826 standard; peptide; 31 AA.
 XX ABU60826;
 XX 06-MAY-2003 (first entry)
 XX Peptide production by gene recombination associated peptide #10.
 XX Peptide production; low-molecular peptide; Kiss-1; GPR8 ligand;
 XX gene recombination.
 XX Rattus sp.
 XX WO200292829-A1.
 XX 21-NOV-2002.
 XX 16-MAY-2002; 2002WO-JP004735.
 XX 17-MAY-2001; 2001JP-00147341.
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX Nishimura O, Suenaga M, Ito T, Kitada C;
 XX WPI; 2003-129302/12.
 XX Process for producing peptides e.g. Kiss-1 peptide and GPR8 ligand for
 PT subsequent applications by gene recombination technique through tandem
 PT repeats to provide precursor protein with specific cleavage sites.
 XX Disclosure; Page 58; 87pp; Japanese.
 PS The invention describes a method of producing a peptide comprising the
 CC excision of the N and C-terminals of a target peptide with enzymes or
 CC chemically through the attached cleavage sites repeated by ligation in a
 CC precursor protein. The method is for producing (low-molecular) peptides
 CC e.g. Kiss-1 peptide and GPR8 ligand for subsequent applications by the
 CC gene recombination technique through tandem repeats to provide a

CC precursor protein with specific cleavage sites. With this method, peptide
 CC production can be carried out easily to provide large quantities of the
 CC required peptides. This is the amino acid sequence of a peptide
 CC associated with the peptide production method of the invention
 XX
 SQ Sequence 31 AA;

Query Match 99.4%; Score 168; DB 6; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.5e-17;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTDINPAWYTGIRPVGR 30
 |||||
 Db 1 SRAHQHSMETRTDINPAWYTGIRPVGR 30

RESULT 13

ABU60837
 ID ABU60837 standard; peptide; 31 AA.

AC ABU60837;

DT 06-MAY-2003 (first entry)

XX Peptide production by gene recombination associated peptide #21.

DE Peptide production; low-molecular peptide; KISS-1; GPR8 ligand;
 XX gene recombination.
 KW
 KW
 XX

OS Rattus sp.

XX WO200292829-A1.

PN 21-NOV-2002.

XX 16-MAY-2002; 2002WO-JP004735.

PF 17-MAY-2001; 2001JP-00147341.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Nishimura O, Suenaga M, Ito T, Kitada C;

XX WPI; 2003-129302/12.

XX Process for producing peptides e.g. KISS-1 peptide and GPR8 ligand for
 PT subsequent applications by gene recombination technique through tandem
 PT repeats to provide precursor protein with specific cleavage sites.
 PT

PS Disclosure; Page 64; 87pp; Japanese.

XX The invention describes a method of producing a peptide comprising the
 CC excision of the N and C-terminals of a target peptide with enzymes or
 CC chemically through the attached cleavage sites repeated by ligation in a
 CC precursor protein. The method is for producing (low-molecular) peptides
 CC e.g. KISS-1 peptide and GPR8 ligand for subsequent applications by the
 CC gene recombination technique through tandem repeats to provide a
 CC precursor protein with specific cleavage sites. With this method, peptide
 CC production can be carried out easily to provide large quantities of the
 CC required peptides. This is the amino acid sequence of a peptide
 CC associated with the peptide production method of the invention
 XX
 SQ Sequence 31 AA;

Query Match 99.4%; Score 168; DB 6; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.5e-17;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTDINPAWYTGIRPVGR 30
 |||||
 Db 1 SRAHQHSMETRTDINPAWYTGIRPVGR 30

RESULT 14

AAW31385
 ID AAW31385 standard; peptide; 32 AA.

XX AAW31385;

XX 27-AUG-2003 (revised)

DT 06-APR-1998 (first entry)

XX Rat type G protein-coupled receptor ligand fragment 2.

DE G protein-coupled receptor; ligand binding; pharmaceutical; modulator;
 KW pituitary; central nervous system; pancreas; prophylactic;
 KW therapeutic agent.
 XX

OS Rattus sp.

XX WO9724436-A2.

XX 10-JUL-1997.

XX 26-DEC-1996; 96WO-JP003821.

XX 28-DEC-1995; 95JP-00343371.

PR 15-MAR-1996; 96JP-00059419.

PR 12-AUG-1996; 96JP-00211805.

PR 18-SEP-1996; 96JP-00246573.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;

PI Kitada C;

XX WPI; 1997-363672/33.

DR N-PSDB; AAV02422.

XX Ligand peptide for G protein-coupled receptor - acts by modulating
 PT function in the central nervous system, pancreas and pituitary gland.

PS Claim 2; Page 179; 258pp; English.

XX This sequence represents a peptide fragment from a novel rat type ligand
 CC polypeptide corresponding to amino acid residues 22 to 53 of the sequence
 CC represented in AAW31383 and is used in an assay to monitor ligand binding
 CC to the G protein-coupled receptor protein. Pharmaceutical compositions
 CC containing this ligand may be used as a pituitary function modulator, a
 CC central nervous system modulator or a pancreatic function modulator. This
 CC ligand could have specific applications as a prophylactic or therapeutic
 CC agent for dementia, depression, hyperkinetic syndrome, disturbance of
 CC consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone
 CC secretory disease, hyper- and polyphagia, hyperlipidaemia,
 CC hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,
 CC amyotrophic lateral sclerosis, acute myocardial infarction, infertility,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis and/or oligosaccharia. Assays can also be developed to screen
 CC compounds which are capable of altering the binding activity of the
 CC ligand affecting activation of the G protein-coupled receptor protein.
 CC (Updated on 27-AUG-2003 to correct OS field.)
 XX

SQ Sequence 32 AA;

Query Match 99.4%; Score 168; DB 2; Length 32;
 Best Local Similarity 100.0%; Pred. No. 1.6e-17;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTDINPAWYTGIRPVGR 30
 |||||
 Db 1 SRAHQHSMETRTDINPAWYTGIRPVGR 30

RESULT 15

AAB10356
ID AAB10356 standard; peptide; 32 AA.
XX
AC AAB10356;
XX
DT 24-NOV-2000 (first entry)
XX
DE Rat oxytocin secretion promoting peptide SEQ ID NO: 19.
XX
DE Rat; oxytocin secretion promoter; G protein-coupled receptor protein;
XX treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
XX caesarean section; artificial fertilization; galactostasis; goat; pig;
XX veterinary medicine; milk production.
XX
OS Rattus sp.
XX
XX WO2000038704-A1.
XX
XX 06-JUL-2000.
XX
XX 22-DEC-1999; 99WO-JP007199.
XX
XX 25-DEC-1998; 98JP-00369585.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Matsumoto H, Kitada C, Hinuma S;
XX WPI; 2000-452298/39.
XX
XX Physiologically-active polypeptide recognized as ligand by G protein-
XX coupled receptor protein, for promoting secretion of oxytocin, as drugs
XX for diseases relating to oxytocin secretion and in veterinary medicine.
XX
XX Disclosure; Page 57; 72pp; Japanese.
XX
XX This invention describes a novel oxytocin secretion-regulating agent
XX which contains a ligand peptide or its salt for the G protein-coupled
XX receptor protein. It is useful in the form of drugs for ameliorating,
XX preventing and treating diseases relating to oxytocin secretion e.g. weak
XX pains and atonic bleeding, before and after expulsion of placenta,
XX uterine recovery failure, caesarean section, stoppage of artificial
XX fertilization or galactostasis and is also applicable in veterinary
XX medicine for promoting milk production in cow, goat and pig. This
XX sequence represents a rat peptide which acts as an oxytocin secretion
XX promoter
XX
XX Sequence 32 AA;
XX
XX Query Match 99.4%; Score 168; DB 3; Length 32;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-17; Indels 0; Gaps 0;
XX Matches 30; Conservative 0; Mismatches 0;
XX
XX 1 SRAHQSMETRPDINPAWYTCGIRPVG 30
XX |||||
XX 1 SRAHQSMETRPDINPAWYTCGIRPVG 30
XX |||||

Search completed: August 12, 2004, 14:43:55
Job time : 66.9099 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:49:10 ; Search time 54.7907 Seconds
(without alignments)
177.617 Million cell updates/sec

Title: US-09-700-643A-3
Perfect score: 169
Sequence: 1 SRAHQSMETRTDINPAWYTGRIQVGR 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1292805 seqs, 313927144 residues

Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	168	99.4	31	9 US-09-932-161-14	Sequence 14, Appl
2	168	99.4	31	13 US-10-044-592-4	Sequence 4, Appl
3	168	99.4	31	13 US-10-044-592-5	Sequence 5, Appl
4	168	99.4	31	14 US-10-096-777-14	Sequence 14, Appl
5	168	99.4	70	13 US-10-044-592-90	Sequence 90, Appl
6	168	99.4	82	13 US-10-044-592-1	Sequence 1, Appl
7	168	99.4	86	13 US-10-044-592-96	Sequence 96, Appl
8	168	99.4	91	13 US-10-044-592-94	Sequence 94, Appl
9	157	92.9	31	9 US-09-932-161-13	Sequence 13, Appl
10	157	92.9	31	13 US-10-044-592-39	Sequence 39, Appl
11	157	92.9	31	14 US-10-096-777-13	Sequence 13, Appl
12	157	92.9	32	13 US-10-044-592-40	Sequence 40, Appl
13	157	92.9	33	13 US-10-044-592-41	Sequence 41, Appl
14	157	92.9	98	13 US-10-044-592-28	Sequence 28, Appl
15	157	92.9	98	13 US-10-044-592-38	Sequence 38, Appl

16	157	92.9	98	13	US-10-044-592-82	Sequence 82, Appl
17	157	92.9	98	13	US-10-044-592-84	Sequence 84, Appl
18	157	92.9	98	13	US-10-044-592-86	Sequence 86, Appl
19	157	92.9	98	13	US-10-044-592-88	Sequence 88, Appl
20	152	89.9	29	13	US-10-044-592-26	Sequence 26, Appl
21	143	84.6	31	9	US-09-932-161-15	Sequence 15, Appl
22	143	84.6	31	14	US-10-096-777-15	Sequence 15, Appl
23	143	84.6	87	13	US-10-044-592-92	Sequence 92, Appl
24	133	78.7	25	13	US-10-044-592-78	Sequence 78, Appl
25	110	65.1	20	9	US-09-932-161-17	Sequence 17, Appl
26	110	65.1	20	13	US-10-044-592-6	Sequence 6, Appl
27	110	65.1	20	14	US-10-096-777-17	Sequence 17, Appl
28	105	62.1	19	13	US-10-044-592-27	Sequence 27, Appl
29	105	62.1	20	9	US-09-932-161-16	Sequence 16, Appl
30	105	62.1	20	13	US-10-044-592-42	Sequence 42, Appl
31	105	62.1	20	14	US-10-096-777-16	Sequence 16, Appl
32	105	62.1	21	13	US-10-044-592-43	Sequence 43, Appl
33	105	62.1	22	13	US-10-044-592-44	Sequence 44, Appl
34	99	58.6	20	9	US-09-932-161-18	Sequence 18, Appl
35	99	58.6	20	14	US-10-096-777-18	Sequence 18, Appl
36	91	53.8	40	13	US-10-044-592-80	Sequence 80, Appl
37	62	36.7	209	13	US-10-108-915-30	Sequence 30, Appl
38	61	36.1	1245	15	US-10-369-493-20447	Sequence 20447, A
39	60	35.5	105	12	US-10-424-599-160071	Sequence 160071, A
40	57	33.7	9	13	US-10-044-592-8	Sequence 8, Appl
41	57	33.7	133	12	US-10-424-599-153474	Sequence 153474, A
42	57	33.7	1003	12	US-10-424-599-160069	Sequence 160069, A
43	56	33.1	428	9	US-09-820-155-2	Sequence 2, Appl
44	53.5	31.7	413	12	US-10-282-122A-53922	Sequence 53922, A
45	53	31.4	465	14	US-10-301-822-197	Sequence 197, App

ALIGNMENTS

RESULT 1
US-09-932-161-14
; Sequence 14, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Civallo, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; FILE OF INVENTION: Promoting Wakefulness and Sleep
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932.161
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Rattus
US-09-932-161-14

Query Match 99.4%; Score 168; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. No. 1e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQSMETRTDINPAWYTGRIQVGR 30
Db 1 SRAHQSMETRTDINPAWYTGRIQVGR 30

RESULT 2
US-10-044-592-4
; Sequence 4, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji

; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 4
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Murine
US-10-044-592-4

Query Match 99.4%; Score 168; DB 13; Length 31;
Best Local Similarity 100.0%; Pred. No. 1e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQHSMETRTDINPAWYTGIRPVGR 30
Db 1 SRAHQHSMETRTDINPAWYTGIRPVGR 30

RESULT 3
US-10-044-592-5
; Sequence 5, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hiruma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 5
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(31)
; OTHER INFORMATION: antigen
US-10-044-592-5

Query Match 99.4%; Score 168; DB 13; Length 31;
Best Local Similarity 100.0%; Pred. No. 1e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQHSMETRTDINPAWYTGIRPVGR 30
Db 1 SRAHQHSMETRTDINPAWYTGIRPVGR 30

RESULT 4
US-10-096-777-14
; Sequence 14, Application US/10096777
; Publication No. US20030171270A1
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Therapeutic Compositions and Methods

; TITLE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Rattus
US-10-096-777-14

Query Match 99.4%; Score 168; DB 14; Length 31;
Best Local Similarity 100.0%; Pred. No. 1e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQHSMETRTDINPAWYTGIRPVGR 30
Db 1 SRAHQHSMETRTDINPAWYTGIRPVGR 30

RESULT 5
US-10-044-592-90
; Sequence 90, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hiruma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 90
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Rat
US-10-044-592-90

Query Match 99.4%; Score 168; DB 13; Length 70;
Best Local Similarity 100.0%; Pred. No. 2.5e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRAHQHSMETRTDINPAWYTGIRPVGR 30
Db 22 SRAHQHSMETRTDINPAWYTGIRPVGR 51

RESULT 6
US-10-044-592-1
; Sequence 1, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hiruma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974

```
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 1
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Murine
US-10-044-592-1

Query Match          99.4%; Score 168; DB 13; Length 82;
Best Local Similarity 100.0%; Pred. No. 2.9e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTDPINPAWYTGIRPVGR 30
    |||||||
Db 21 SRAHQHSMETRTDPINPAWYTGIRPVGR 50
    |||||||

RESULT 7
US-10-044-592-96
; Sequence 96, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 96
; LENGTH: 86
; TYPE: PRT
; ORGANISM: mammalian
; NAME/KEY: misc feature
; LOCATION: (1224)..(1243)
; OTHER INFORMATION: Bracket region depicted in FIG 39.
US-10-044-592-96

Query Match          99.4%; Score 168; DB 13; Length 86;
Best Local Similarity 100.0%; Pred. No. 3.1e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTDPINPAWYTGIRPVGR 30
    |||||||
Db 22 SRAHQHSMETRTDPINPAWYTGIRPVGR 51
    |||||||

RESULT 8
US-10-044-592-94
; Sequence 94, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28

; PRIOR FILING DATE: 1997-04-28

; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 94
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Mammalian
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(31)
; OTHER INFORMATION: Primer
; NAME/KEY: misc feature
; LOCATION: (925)..(955)
; OTHER INFORMATION: primer
; NAME/KEY: misc feature
; LOCATION: (1)..(955)
; OTHER INFORMATION: Insert fragment of pmGB3 as depicted in FIG 36 and 37.
US-10-044-592-94

Query Match          99.4%; Score 168; DB 13; Length 91;
Best Local Similarity 100.0%; Pred. No. 3.2e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTDPINPAWYTGIRPVGR 30
    |||||||
Db 22 SRAHQHSMETRTDPINPAWYTGIRPVGR 51
    |||||||

RESULT 9
US-09-932-161-13
; Sequence 13, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; TITLE OF INVENTION: Promoting Wakefulness and Sleep
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-932-161-13

Query Match          92.9%; Score 157; DB 9; Length 31;
Best Local Similarity 93.3%; Pred. No. 3.9e-15;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTDPINPAWYTGIRPVGR 30
    |||||||
Db 1 SRAHQHSMETRTDPINPAWYTGIRPVGR 30
    |||||||

RESULT 10
US-10-044-592-39
; Sequence 39, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
```

```

; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 39
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-39

Query Match          92.9%; Score 157; DB 13; Length 31;
Best Local Similarity 93.3%; Pred. No. 3.9e-15;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SRAHQSMETRTDINPANYTGRGIRPVGR 30
Db 1 SRAHQSMETRTDINPANYAGRGIRPVGR 30

RESULT 11
US-10-096-777-13
; Sequence 13, Application US/10096777
; Publication No. US20030171270A1
; GENERAL INFORMATION:
; APPLICANT: Civelii, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; TITLE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)
; FILE REFERENCE: PUC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Bos taurus
US-10-096-777-13

Query Match          92.9%; Score 157; DB 14; Length 31;
Best Local Similarity 93.3%; Pred. No. 3.9e-15;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SRAHQSMETRTDINPANYTGRGIRPVGR 30
Db 1 SRAHQSMETRTDINPANYAGRGIRPVGR 30

RESULT 12
US-10-044-592-40
; Sequence 40, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 40
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-40

Query Match          92.9%; Score 157; DB 13; Length 32;
Best Local Similarity 93.3%; Pred. No. 1.3e-14;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SRAHQSMETRTDINPANYTGRGIRPVGR 30
Db 1 SRAHQSMETRTDINPANYAGRGIRPVGR 30

RESULT 13
US-10-044-592-41
; Sequence 41, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 41
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-41

Query Match          92.9%; Score 157; DB 13; Length 33;
Best Local Similarity 93.3%; Pred. No. 4.1e-15;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SRAHQSMETRTDINPANYTGRGIRPVGR 30
Db 1 SRAHQSMETRTDINPANYAGRGIRPVGR 30

RESULT 14
US-10-044-592-28
; Sequence 28, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 28
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Murine
US-10-044-592-28

Query Match          92.9%; Score 157; DB 13; Length 98;
Best Local Similarity 93.3%; Pred. No. 1.3e-14;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

QY 1 SRAHQHSMETRTDINPAWYTGRIQVGR 30
 |||||
 Db 23 SRAHQHSMETRTDINPAWYAGRIQVGR 52
 |||||

RESULT 15

US-10-044-592-38
 ; Sequence 38, Application US/10044592
 ; Publication No. US20020143152A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hinuma, Shuji
 ; APPLICANT: Fukusumi, Shoji
 ; TITLE OF INVENTION: Polypeptides, their Production and Use
 ; FILE REFERENCE: 2463US2P
 ; CURRENT APPLICATION NUMBER: US/10/044,592
 ; CURRENT FILING DATE: 2002-01-10
 ; PRIOR APPLICATION NUMBER: US 09/403639
 ; PRIOR FILING DATE: 1999-25-10
 ; PRIOR APPLICATION NUMBER: PCT/JP98/01923
 ; PRIOR FILING DATE: 1998-04-27
 ; PRIOR APPLICATION NUMBER: JP 9-109974
 ; PRIOR FILING DATE: 1997-04-28
 ; NUMBER OF SEQ.ID NOS: 96
 ; SOFTWARE:
 ; SEQ ID NO 38
 ; LENGTH: 98
 ; TYPE: PRT
 ; ORGANISM: Bovine
 US-10-044-592-38

Query Match 92.9%; Score 157; DB 13; Length 98;
 Best Local Similarity 93.3%; Pred. No. 1.3e-14;
 Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTDINPAWYTGRIQVGR 30
 |||||
 Db 23 SRAHQHSMETRTDINPAWYAGRIQVGR 52
 |||||

Search completed: August 12, 2004, 15:22:50
 Job time : 54.7907 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:36 ; Search time 17.6628 Seconds
(without alignments)
90.609 Million cell updates/sec

Title: US-09-700-643A-3
Perfect score: 169
Sequence: 1 SRAHQSMETRPDINPAWYTGIRPVGRX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/prodata/2/iaa/5A COMB.pcp.*
- 2: /cgn2_6/prodata/2/iaa/5B COMB.pcp.*
- 3: /cgn2_6/prodata/2/iaa/6A COMB.pcp.*
- 4: /cgn2_6/prodata/2/iaa/6B COMB.pcp.*
- 5: /cgn2_6/prodata/2/iaa/PCTUS COMB.pcp.*
- 6: /cgn2_6/prodata/2/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	168	99.4	31	US-09-105-678A-8	Sequence 8, Appl
2	168	99.4	31	US-09-105-678A-37	Sequence 37, Appl
3	168	99.4	31	US-09-105-678A-37	Sequence 4, Appl
4	168	99.4	31	US-09-172-353-4	Sequence 47, Appl
5	168	99.4	31	US-08-776-971-47	Sequence 8, Appl
6	168	99.4	31	US-09-421-208-8	Sequence 37, Appl
7	168	99.4	31	US-09-421-208-37	Sequence 14, Appl
8	168	99.4	31	US-09-560-915-14	Sequence 4, Appl
9	168	99.4	32	US-09-799-955-4	Sequence 38, Appl
10	168	99.4	32	US-08-776-971-48	Sequence 48, Appl
11	168	99.4	32	US-09-421-208-38	Sequence 38, Appl
12	168	99.4	33	US-09-105-678A-39	Sequence 39, Appl
13	168	99.4	33	US-08-776-971-49	Sequence 49, Appl
14	168	99.4	33	US-09-421-208-39	Sequence 39, Appl
15	168	99.4	83	US-08-776-971-45	Sequence 45, Appl
16	168	99.4	83	US-08-776-971-124	Sequence 124, App
17	168	99.4	83	US-08-776-971-137	Sequence 137, App
18	157	92.9	21	US-09-105-678A-7	Sequence 7, Appl
19	157	92.9	31	US-09-105-678A-31	Sequence 31, Appl
20	157	92.9	31	US-08-776-971-5	Sequence 5, Appl
21	157	92.9	31	US-08-776-971-97	Sequence 97, Appl
22	157	92.9	31	US-09-421-208-7	Sequence 7, Appl
23	157	92.9	31	US-09-421-208-31	Sequence 31, Appl
24	157	92.9	31	US-09-560-915-13	Sequence 13, Appl
25	157	92.9	32	US-09-105-678A-32	Sequence 32, Appl
26	157	92.9	32	US-08-776-971-6	Sequence 6, Appl
27	157	92.9	32	US-09-421-208-32	Sequence 32, Appl

28	157	92.9	33	US-09-105-678A-33	Sequence 33, Appl
29	157	92.9	33	US-08-776-971-7	Sequence 7, Appl
30	157	92.9	33	US-09-421-208-33	Sequence 33, Appl
31	157	92.9	98	US-08-776-971-1	Sequence 1, Appl
32	157	92.9	98	US-08-776-971-44	Sequence 44, Appl
33	157	92.9	98	US-08-776-971-115	Sequence 115, App
34	157	92.9	98	US-08-776-971-117	Sequence 117, App
35	157	92.9	98	US-08-776-971-122	Sequence 122, App
36	157	92.9	98	US-08-776-971-131	Sequence 131, App
37	157	92.9	98	US-08-776-971-136	Sequence 136, App
38	152	89.9	29	US-09-105-678A-29	Sequence 29, Appl
39	152	89.9	29	US-08-776-971-3	Sequence 3, Appl
40	152	89.9	29	US-09-421-208-29	Sequence 29, Appl
41	143	84.6	31	US-09-105-678A-9	Sequence 43, Appl
42	143	84.6	31	US-09-105-678A-43	Sequence 61, Appl
43	143	84.6	31	US-08-776-971-61	Sequence 9, Appl
44	143	84.6	31	US-09-421-208-9	Sequence 43, Appl
45	143	84.6	31	US-09-421-208-43	Sequence 43, Appl

ALIGNMENTS

RESULT 1
US-09-105-678A-8
; Sequence 8, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-105-678A-8

Query Match 99.4%; Score 168; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 9e-19;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQSMETRPDINPAWYTGIRPVGR 30

Db 1 SRAHQSMETRTDINPAWYTGIRPVGR 30
|||||
RESULT 2
US-09-105-678A-37
; Sequence 37, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-105-678A-37
Query Match 99.4%; Score 168; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 9e-19;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQSMETRTDINPAWYTGIRPVGR 30
|||||
Db 1 SRAHQSMETRTDINPAWYTGIRPVGR 30
|||||
RESULT 3
US-09-172-353-4
; Sequence 4, Application US/09172353
; Patent No. 6197530
; GENERAL INFORMATION:
; APPLICANT: Stricker-Kongra, Alain
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: GPRI0 AS A TARGET FOR IDENTIFYING WEIGHT MODULATING COMPOUNDS
; FILE REFERENCE: 07334/102001
; CURRENT APPLICATION NUMBER: US/09/172,353
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 31

; TYPE: PPT
; ORGANISM: Mus musculus
US-09-172-353-4
Query Match 99.4%; Score 168; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 9e-19;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQSMETRTDINPAWYTGIRPVGR 30
|||||
Db 1 SRAHQSMETRTDINPAWYTGIRPVGR 30
|||||
RESULT 4
US-08-776-971-47
; Sequence 47, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; ADDRESSEE: Habata, Yugo
; ADDRESSEE: Kawamata, Yuji
; ADDRESSEE: Hosoya, Masaki
; ADDRESSEE: Fujii, Ryo
; ADDRESSEE: Fukusumi, Shoji
; ADDRESSEE: Kitada, Chieko
; TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 47:
US-08-776-971-47

Query Match 99.4%; Score 168; DB 3; Length 31;

APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu

QY 1 SRAHQSMETRTPDINPAWYTCRGIRPVGR 30

Db 1 SRAHQSMETRTDINPAWYTGIRPVGR 30

RESULT 8

US-09-799-955-4

Sequence 4, Application US/09799955

Patent No. 6537765

GENERAL INFORMATION:

APPLICANT: Stricker-Kongra, Alain

APPLICANT: Gu, Wei

TITLE OF INVENTION: GPR10 AS A TARGET FOR IDENTIFYING WEIGHT MODULATING COMPOUNDS

FILE REFERENCE: 07334/102001

CURRENT APPLICATION NUMBER: US/09/799,955

CURRENT FILING DATE: 2001-03-06

PRIOR APPLICATION NUMBER: US/09/172,353

PRIOR FILING DATE: 1998-10-14

NUMBER OF SEQ ID NOS: 7

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 4

LENGTH: 31

TYPE: PRT

ORGANISM: Mus musculus

US-09-799-955-4

Query Match 99.4%; Score 168; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 9e-19;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQSMETRTDINPAWYTGIRPVGR 30

Db 1 SRAHQSMETRTDINPAWYTGIRPVGR 30

RESULT 9

US-09-105-678A-38

Sequence 38, Application US/09105678A

Patent No. 6103682

GENERAL INFORMATION:

APPLICANT: Suenaga, Masato

APPLICANT: Moriya, Takeo

APPLICANT: Tanaka, Yoko

APPLICANT: Nishimura, Osamu

TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/105,678A

FILING DATE: 26-JUN-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 172118/1997

FILING DATE: 27-JUN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 48466-342

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 38:

SEQUENCE CHARACTERISTICS:

LENGTH: 32 amino acids

TYPE: amino acid

Db 1 SRAHQSMETRTDINPAWYTGIRPVGR 30

Query Match 99.4%; Score 168; DB 3; Length 32;

Best Local Similarity 100.0%; Pred. No. 9.4e-19;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQSMETRTDINPAWYTGIRPVGR 30

Db 1 SRAHQSMETRTDINPAWYTGIRPVGR 30

RESULT 10

US-08-776-971-48

Sequence 48, Application US/08776971B

Patent No. 6228984

GENERAL INFORMATION:

APPLICANT: Hiruma, Shuji

APPLICANT: Habata, Yugo

APPLICANT: Kawamata, Yuji

APPLICANT: Hosoya, Masaki

APPLICANT: Fujii, Ryo

APPLICANT: Fukusumi, Shoji

APPLICANT: Kitada, Chieko

TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE

NUMBER OF SEQUENCES: 140

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/776,971B

FILING DATE: 06-Feb-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP96/03821

FILING DATE: 28-DEC-1996

APPLICATION NUMBER: JP 7/343371

FILING DATE: 28-DEC-1995

APPLICATION NUMBER: JP 8/59419

FILING DATE: 15-MAR-1996

APPLICATION NUMBER: JP 8/211805

FILING DATE: 12-AUG-1996

APPLICATION NUMBER: JP 8/246573

FILING DATE: 18-SEP-1996

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 47176

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 48:

SEQUENCE CHARACTERISTICS:

LENGTH: 32 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: internal

SEQUENCE DESCRIPTION: SEQ ID NO: 48:

US-08-776-971-48

Query Match 99.4%; Score 168; DB 3; Length 32;
 Best Local Similarity 100.0%; Pred. No. 9.4e-19;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTPDINPAWYTGIRPVGR 30
 DB 1 SRAHQHSMETRTPDINPAWYTGIRPVGR 30

RESULT 11
 US-09-421-208-38
 ; Sequence 38, Application US/09421208
 ; Patent No. 6258561
 ; GENERAL INFORMATION:
 ; APPLICANT: Suenaga, Masato
 ; APPLICANT: Moriya, Takeo
 ; APPLICANT: Tanaka, Yoko
 ; APPLICANT: Nishimura, Osamu
 ; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
 ; NUMBER OF SEQUENCES: 52
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
 ; STREET: 130 Water Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/421,208
 ; FILING DATE:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 09/105,678
 ; FILING DATE: 26-JUN-1998
 ; APPLICATION NUMBER: JP 172118/1997
 ; FILING DATE: 27-JUN-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Conlin, David G.
 ; REGISTRATION NUMBER: 27,026
 ; REFERENCE/DOCKET NUMBER: 48466-342
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-523-3400
 ; TELEFAX: 617-523-6440
 ; INFORMATION FOR SEQ ID NO: 38:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 32 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-09-421-208-38

Query Match 99.4%; Score 168; DB 3; Length 32;
 Best Local Similarity 100.0%; Pred. No. 9.4e-19;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTPDINPAWYTGIRPVGR 30
 DB 1 SRAHQHSMETRTPDINPAWYTGIRPVGR 30

RESULT 12
 US-09-105-678A-39
 ; Sequence 39, Application US/09105678A
 ; Patent No. 6103882
 ; GENERAL INFORMATION:
 ; APPLICANT: Suenaga, Masato
 ; APPLICANT: Moriya, Takeo
 ; APPLICANT: Tanaka, Yoko

Query Match 99.4%; Score 168; DB 3; Length 32;
 Best Local Similarity 100.0%; Pred. No. 9.4e-19;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTPDINPAWYTGIRPVGR 30
 DB 1 SRAHQHSMETRTPDINPAWYTGIRPVGR 30

RESULT 12
 US-09-105-678A-39
 ; Sequence 39, Application US/09105678A
 ; Patent No. 6103882
 ; GENERAL INFORMATION:
 ; APPLICANT: Suenaga, Masato
 ; APPLICANT: Moriya, Takeo
 ; APPLICANT: Tanaka, Yoko

Query Match 99.4%; Score 168; DB 3; Length 33;
 Best Local Similarity 100.0%; Pred. No. 9.7e-19;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTPDINPAWYTGIRPVGR 30
 DB 1 SRAHQHSMETRTPDINPAWYTGIRPVGR 30

RESULT 11
 US-09-421-208-38
 ; Sequence 38, Application US/09421208
 ; Patent No. 6258561
 ; GENERAL INFORMATION:
 ; APPLICANT: Suenaga, Masato
 ; APPLICANT: Moriya, Takeo
 ; APPLICANT: Tanaka, Yoko
 ; APPLICANT: Nishimura, Osamu
 ; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
 ; NUMBER OF SEQUENCES: 52
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
 ; STREET: 130 Water Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/105,678A
 ; FILING DATE: 26-JUN-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 172118/1997
 ; FILING DATE: 27-JUN-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Conlin, David G.
 ; REGISTRATION NUMBER: 27,026
 ; REFERENCE/DOCKET NUMBER: 48466-342
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-523-3400
 ; TELEFAX: 617-523-6440
 ; INFORMATION FOR SEQ ID NO: 39:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 33 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-09-105-678A-39

Query Match 99.4%; Score 168; DB 3; Length 33;
 Best Local Similarity 100.0%; Pred. No. 9.7e-19;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTPDINPAWYTGIRPVGR 30
 DB 1 SRAHQHSMETRTPDINPAWYTGIRPVGR 30

RESULT 13
 US-08-776-971-49
 ; Sequence 49, Application US/08776971B
 ; Patent No. 6228984
 ; GENERAL INFORMATION:
 ; APPLICANT: Hinuma, Shuji
 ; APPLICANT: Habata, Yugo
 ; APPLICANT: Kawamata, Yuji
 ; APPLICANT: Hosoya, Masaki
 ; APPLICANT: Fujii, Ryo
 ; APPLICANT: Fukusumi, Shoji
 ; APPLICANT: Kitada, Chieko
 ; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
 ; NUMBER OF SEQUENCES: 140
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
 ; STREET: 130 Water Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:

Query Match 99.4%; Score 168; DB 3; Length 33;
 Best Local Similarity 100.0%; Pred. No. 9.7e-19;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQHSMETRTPDINPAWYTGIRPVGR 30
 DB 1 SRAHQHSMETRTPDINPAWYTGIRPVGR 30

RESULT 13
 US-08-776-971-49
 ; Sequence 49, Application US/08776971B
 ; Patent No. 6228984
 ; GENERAL INFORMATION:
 ; APPLICANT: Hinuma, Shuji
 ; APPLICANT: Habata, Yugo
 ; APPLICANT: Kawamata, Yuji
 ; APPLICANT: Hosoya, Masaki
 ; APPLICANT: Fujii, Ryo
 ; APPLICANT: Fukusumi, Shoji
 ; APPLICANT: Kitada, Chieko
 ; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
 ; NUMBER OF SEQUENCES: 140
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
 ; STREET: 130 Water Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/776,971B
FILING DATE: 06-FEB-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03821
FILING DATE: 28-DEC-1996
APPLICATION NUMBER: JP 7/343371
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: JP 8/59419
FILING DATE: 15-MAR-1996
APPLICATION NUMBER: JP 8/211805
FILING DATE: 12-AUG-1996
APPLICATION NUMBER: JP 8/246573
FILING DATE: 18-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 47176
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 49:
US-08-776-971-49
Query Match 99.4%; Score 168; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 9,7e-19;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQSMETRTDINPAWYTGRIQVGR 30
Db 1 SRAHQSMETRTDINPAWYTGRIQVGR 30
RESULT 14
US-09-421-208-39
Sequence 39, Application US/09421208
Patent No. 6258561
GENERAL INFORMATION:
APPLICANT: Sueraga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/421,208
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/105,678
FILING DATE: 26-JUN-1998
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-421-208-39
Query Match 99.4%; Score 168; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 9,7e-19;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHQSMETRTDINPAWYTGRIQVGR 30
Db 1 SRAHQSMETRTDINPAWYTGRIQVGR 30
RESULT 15
US-08-776-971-45
Sequence 45, Application US/08776971B
Patent No. 6228984
GENERAL INFORMATION:
APPLICANT: Hinuma, Shuji
APPLICANT: Habata, Yugo
APPLICANT: Kawamata, Yuji
APPLICANT: Hosoya, Masaki
APPLICANT: Fujii, Ryo
APPLICANT: Fukusumi, Shoji
APPLICANT: Kitada, Chieko
TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,971B
FILING DATE: 06-FEB-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03821
FILING DATE: 28-DEC-1996
APPLICATION NUMBER: JP 7/343371
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: JP 8/59419
FILING DATE: 15-MAR-1996
APPLICATION NUMBER: JP 8/211805
FILING DATE: 12-AUG-1996
APPLICATION NUMBER: JP 8/246573
FILING DATE: 18-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 47176
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 45:

```
;
;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 83 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   FRAGMENT TYPE: internal
;   SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-08-776-971-45

Query Match          99.4%; Score 168; DB 3; Length 83;
Best Local Similarity 100.0%; Pred. No. 2.9e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SRAHQHSMETRPDINPAWYTGIRPVGR 30
        |||||
Db       22 SRAHQHSMETRPDINPAWYTGIRPVGR 51

Search completed: August 12, 2004, 14:52:13
Job time : 17.6628 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:36 ; Search time 8.83721 Seconds
(without alignments)
217.697 Million cell updates/sec

Title: US-09-700-643A-5
Perfect score: 108
Sequence: 1 TPDINPAWYASRGIRPVGRX 20

Scoring table: BLOSUM62
Gapop 10.0 , Capext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: Piri:*
2: Pir2:*
3: Pir3:*
4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	91.7	83	JC7607	prolactin-releasing
2	53	49.1	790	T47959	hypothetical prote
3	50.5	46.8	664	F83376	conserved hypothet
4	47	43.5	284	F71015	hypothetical prote
5	47	43.5	333	H82852	hydroxybenzoate oc
6	47	43.5	938	C84480	hypothetical prote
7	46.5	43.1	779	T49717	related to BCS1 pr
8	46	42.6	419	AH3166	hypothetical prote
9	45	41.7	390	G82844	cysteine synthase
10	45	41.7	430	B69009	conserved hypothet
11	45	41.7	787	T21869	hypothetical prote
12	44	40.7	276	D70817	hypothetical prote
13	44	40.7	284	A75117	hypothetical prote
14	44	40.7	302	B70631	hypothetical prote
15	44	40.7	309	T32376	hypothetical prote
16	44	40.7	428	F81694	pyruvate dehydroge
17	44	40.7	476	G64720	probable amino aci
18	44	40.7	476	G90829	probable inner mem
19	44	40.7	476	G85480	inner membrane tra
20	44	40.7	527	T33175	hypothetical prote
21	44	40.7	719	S61046	ARPI protein - yea
22	43.5	40.3	107	D87337	hypothetical prote
23	43.5	40.3	374	G70947	hypothetical prote
24	43	39.8	181	AE2305	hypothetical prote
25	43	39.8	220	C83292	probable glutathio
26	43	39.8	250	G83400	hypothetical prote
27	43	39.8	376	A48197	opsin, ocular - A
28	43	39.8	376	B48197	opsin, lateral eye
29	43	39.8	433	H87660	peptidoglycan-bind

30	43	39.8	468	2	C83160	nitrite extrusion
31	43	39.8	1359	2	T10235	xanthine dehydroge
32	43	39.8	1364	2	T10236	xanthine dehydroge
33	42.5	39.4	345	2	D84012	N-acetylglutamate
34	42.5	39.4	3944	2	T19997	hypothetical prote
35	42	38.9	105	2	A97505	hypothetical prote
36	42	38.9	105	2	AE2723	hypothetical prote
37	42	38.9	113	2	AI2484	hypothetical prote
38	42	38.9	128	2	S76955	hypothetical prote
39	42	38.9	132	2	B82824	protein-export mem
40	42	38.9	216	2	H82608	transcription regu
41	42	38.9	219	2	H84594	hypothetical prote
42	42	38.9	221	2	A75366	probable transcrip
43	42	38.9	232	2	C75608	hypothetical prote
44	42	38.9	311	2	S86600	cytochrome-c oxida
45	42	38.9	335	2	B87151	LytB homolog [imp

ALIGNMENTS

RESULT 1

JC7607
prolactin-releasing peptide - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C:Accession: JC7607
Riyamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Hosoya, T
Biochem. Biophys. Res. Commun. 281, 53-56, 2001
A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene: Mul
A:Reference number: JC7607; MUID:21092785; PMID:11178959
A:Contents: Spleen
A:Accession: JC7607
A:Molecule type: DNA
A:Residues: 1-83 <YAM>
A:Cross-references: DBJ:AB040612; DBJ:AB040613
C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior pit
release, and stimulation of ACTH secretion from the pituitary.
C:Genetics:
A:Gene: PrRP
A:Introns: 33/1

Query Match 91.7%; Score 99; DB 2; Length 83;
Best Local Similarity 89.5%; Pred. No. 2.3e-08;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGR 19
Db 33 TPDINPAWYASRGIRPVGR 51

RESULT 2

T47959
hypothetical protein F15G16.60 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47959
R:De Haan, M.; Maarse, A.C.; Grivell, L.A.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quet
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z24480
A:Accession: T47959
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-790 <DEH>
A:Cross-references: EMBL:AL132959
A:Experimental source: cultivar Columbia; BAC clone F15G16
C:Genetics:
A:Map position: 3
A:Introns: 39/1; 678/2; 698/3; 773/2
A>Note: F15G16.60

Query Match 49.1%; Score 53; DB 2; Length 790;
Best Local Similarity 55.6%; Pred. No. 3.4;

Accession: E000000000
A>Status: preliminary; translated from GB/EMBL/DBJ
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-767 <WIL>

A;Cross-references: EMBL:Z68342; PIDN:CAA92775.1; GSPDB:GN00022; CESP:F38E11.7
 A;Experimental source: clone F38E11
 C;Genetics:
 A;Map position: 4
 A;Introns: 50/2; 118/1; 139/2; 189/3; 226/1; 248/1; 287/2; 375/2; 432/3; 465/3; 548/1; 6

Query Match 41.7%; Score 45; DB 2; Length 767;
 Best Local Similarity 75.0%; Pred. No. 58;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PDINPAWY 9
 ||: |||||

Db 747 PDVKPAWY 754

RESULT 12
 Hypothetical protein Rv1716 - Mycobacterium tuberculosis (strain H37RV)
 C;Species: Mycobacterium tuberculosis
 C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C;Accession: D70817
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A;Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A;Reference number: A70500; MUID:98295987; PMID:9634230
 A;Accession: D70817
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-276 <COL>
 A;Cross-references: GB:AL022003; GB:AL123456; NID:G3261547; PIDN:CAA17613.1; PID:G326154
 A;Experimental source: strain H37RV
 C;Genetics:
 A;Gene: Rv1716

Query Match 40.7%; Score 44; DB 2; Length 276;
 Best Local Similarity 20.0%; Pred. No. 29;
 Matches 10; Conservative 5; Mismatches 3; Indels 32; Gaps 1;

Qy 1 TPDINPA-----WYASRGIRPVG 18
 |||||

Db 111 TPIRFGDIVVNTGWHKKAUSAEYIAYSPGFKKAGWFAKGVKAVG 160

RESULT 13
 Hypothetical protein PAB0502 - Pyrococcus abyssi (strain Orsay)
 C;Species: Pyrococcus abyssi
 C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 C;Accession: A75117
 R;anonymous, Genoscope
 A;Submitted to the EMBL Data Library, July 1999
 A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
 A;Reference number: A75001
 A;Accession: A75117
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-284 <XAW>
 A;Cross-references: GB:AJ248285; GB:AL096836; NID:G5458067; PIDN:CAB49650.1; PID:G545816
 A;Experimental source: strain Orsay
 C;Genetics:
 A;Gene: PAB0502
 C;Superfamily: Pyrococcus horikoshii hypothetical protein PH1420

Query Match 40.7%; Score 44; DB 2; Length 284;
 Best Local Similarity 44.4%; Pred. No. 30;
 Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 2 PDINPAWYASRGIRPVGR 19
 ||||| :||| :||| :|||

Db 217 PYIBPTFVALRGLLELCK 234

RESULT 14
 B70631

Hypothetical protein Rv0428c - Mycobacterium tuberculosis (strain H37RV)

C;Species: Mycobacterium tuberculosis
 C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C;Accession: B70631

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A;Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A;Reference number: A70500; MUID:98295987; PMID:9634230
 A;Accession: B70631
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-302 <COL>
 A;Cross-references: GB:Z84724; GB:AL123456; NID:G3261708; PIDN:CAB06568.1; PID:G1817694
 A;Experimental source: strain H37RV
 C;Genetics:
 A;Gene: Rv0428c
 C;Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0428c

Query Match 40.7%; Score 44; DB 2; Length 302;
 Best Local Similarity 70.0%; Pred. No. 32;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 7 AWYASRGIRP 16
 ||||| :|||

Db 126 AWYASRDLPQ 135

RESULT 15
 T32376

Hypothetical protein K10F12.4 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans
 C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C;Accession: T32376
 R;Wohlmann, P.; Beck, C.
 submitted to the EMBL Data Library, September 1997
 A;Description: The sequence of C. elegans cosmid K10F12.
 A;Reference number: Z21157
 A;Accession: T32376
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-309 <WOH>
 A;Cross-references: EMBL:AF025462; PIDN:AAB71002.1; GSPDB:GN00021; CESP:K10F12.4
 A;Experimental source: strain Bristol N2, clone K10F12
 C;Genetics:
 A;Gene: CESP:K10F12.4
 A;Map position: 3
 A;Introns: 31/3; 123/2; 196/3; 239/1

Query Match 40.7%; Score 44; DB 2; Length 309;
 Best Local Similarity 44.4%; Pred. No. 33;
 Matches 8; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

Qy 2 PDINPAWYASRGIRPVGR 19
 ||||| :||| :||| :|||

Db 131 PDRSPWYLPKS--PIGR 146

Search completed: August 12, 2004, 14:50:28
 Job time : 8.83721 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 5.23256 Seconds
(without alignments)
199.024 Million cell updates/sec

Title: US-09-700-643A-5
Perfect score: 108
Sequence: 1 TPDINPAWYASRGIRPVGRX 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107	99.1	87	1 PRRP_HUMAN	P81277 homo sapien
2	103	95.4	98	1 PRRP_BOVIN	P81264 bos taurus
3	99	91.7	83	1 PRRP_RAT	P81278 rattus norv
4	47	43.5	402	1 EX7L_STRCO	O9f8m3 streptomyc
5	46	42.6	383	1 CYCE_CHRVI	O82947 chromatium
6	46	42.6	413	1 EX7L_CORBP	O8f6p1 corynebacte
7	46	42.6	417	1 EX7L_CORGL	O8nm3 corynebacte
8	44	40.7	476	1 YAAJ_ECOLI	P30143 escherichia
9	44	40.7	719	1 NRPI_YEAST	P32770 saccharomyc
10	43.5	40.3	860	1 VG12_BPB03	C37893 bacterioph
11	43	39.8	211	1 GIDB_PSESM	O87ts4 pseudomonas
12	43	39.8	376	1 OPSI_LIMPO	P35360 limulus pol
13	43	39.8	376	1 OPS2_LIMPO	P35361 limulus pol
14	43	39.8	838	1 GLGB_STRAW	O82jfo streptomyc
15	42.5	39.4	345	1 ARGC_BACHD	O9k8v2 bacillus ha
16	42	38.9	332	1 ISPH_MYCLE	O9x781 mycobacteri
17	42	38.9	364	1 YAIW_ECOLI	P77562 escherichia
18	42	38.9	459	1 HMT2_SCHPO	O94284 schizosacch
19	42	38.9	512	1 UBP3_SCHPO	O94269 schizosacch
20	42	38.9	581	1 POL_MLVKK	P31795 radiation m
21	42	38.9	843	1 POL_MLVAK	P03357 akr murine
22	42	38.9	1024	1 POPC_RALSO	O93b52 ralstonia s
23	42	38.9	1196	1 POL_MLVAV	P03356 akr murine
24	42	38.9	1196	1 POL_MLVVD	P11227 radiation m
25	41.5	38.4	1400	1 RIFI_SCHPO	O96up3 schizosacch
26	41	38.0	342	1 Y762_METJA	O58172 methanococc
27	41	38.0	347	1 Y576_METJA	P29978 methanococc
28	41	38.0	384	1 YE48_ANASP	P29978 anabaena sp
29	41	38.0	622	1 PPCC_RAT	P07379 rattus norv
30	41	38.0	676	1 EXLI_HUMAN	O92935 homo sapien
31	41	38.0	697	1 LCFF_HUMAN	O9uku0 homo sapien
32	41	38.0	986	1 CYGF_ARBPV	P11528 arabacia pun
33	40	37.0	155	1 HEXE_ADEB2	O89592 bovine aden

34	40	37.0	158	1 RL15_AERPE	Q9yf98 aeropyrum p
35	40	37.0	208	1 YK69_CABEL	P34347 caenorhabdi
36	40	37.0	328	1 CNA4_HUMAN	O9nvr5 homo sapien
37	40	37.0	424	1 SAVB_SALTY	P23832 salmonella
38	40	37.0	556	1 DPOL_WHYW6	P11292 woodchuck h
39	40	37.0	580	1 ATPU_YEAST	P22136 saccharomyc
40	40	37.0	626	1 THIC_PSEPK	Q88da5 pseudomonas
41	40	37.0	757	1 INMT_MOUSE	O8cag8 mus musculu
42	40	37.0	788	1 REC2_HA2IN	P44408 haemophilus
43	40	37.0	792	1 SYFB_XYLFA	O9pfd6 xyella fas
44	40	37.0	792	1 SYFB_XYLFT	O87ab6 xyella fas
45	40	37.0	879	1 DPOL_WHV1	P03160 woodchuck h

ALIGNMENTS

RESULT 1					
PRRP_HUMAN					
ID	PRRP_HUMAN	STANDARD;	PRT;	87 AA.	
AC	P81277;				
DT	30-MAY-2000 (Rel. 39, Created)				
DT	30-MAY-2000 (Rel. 39, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing				
DE	hormone) (Contains: Prolactin-releasing peptide PrRP31; Prolactin-				
DE	releasing peptide PrRP20].				
GN	PRH.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Brain;				
RX	MEDLINE=98268781; PubMed=9607765;				
RA	Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,				
RA	Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,				
RA	Kurokawa T., Nishimura O., Onda H., Fujino M.,				
RT	"A prolactin-releasing peptide in the brain.";				
RL	Nature 393:272-276(1998).				
RL	[2]				
TISSUE	SPECIFICITY.				
RX	MEDLINE=99426652; PubMed=10498338;				
RA	Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,				
RA	Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,				
RA	Sumino Y., Fujino M.;				
RT	"Tissue distribution of prolactin-releasing peptide (PrRP) and its				
RT	receptor.";				
RL	Regul. Pept. 83:1-10(1999).				
CC	FUNCTION: Stimulates prolactin (PRL) release and regulates the				
CC	expression of prolactin through its receptor GPR10. May stimulate				
CC	lactotrophs directly to secrete PRL.				
CC	TISSUE SPECIFICITY: Medulla oblongata and hypothalamus.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
CC	EMBL; AB015419; BAA29027.1; ..				
DR	MIM; 602663; ..				
DR	GO; GO:0005180; F:peptide hormone; TAS.				
DR	Hormone; Amidation; Signal.				
FT	SIGNAL	1	22	BY SIMILARITY.	
FT	PEPTIDE	23	53	PROLACTIN-RELEASING PEPTIDE PRRP31.	
FT	PEPTIDE	34	53	PROLACTIN-RELEASING PEPTIDE PRRP20.	
FT	PROPEP	58	87		
FT	MOD RES	53	53	AMIDATION (G-54 PROVIDE AMIDE GROUP).	
FT	SEQUENCE	87 AA;	9639 MW;	229A2F3F50CF981B CRC64;	

Query Match 99.1%; Score 107; DB 1; Length 87;
 Best Local Similarity 100.0%; Pred. No. 2.9e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWASRGIRPVGR 19
 DB 34 TPDINPAWASRGIRPVGR 52

RESULT 2

PRRP BOVIN STANDARD; PRT; 98 AA.

AC P81264; (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing
 hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-
 releasing peptide PrRP20].
 DE hormone)
 DE releasing peptide PrRP20].
 GN PRH.

OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.

RC TISSUE=Brain;
 RX MEDLINE=98268781; PubMed=9607765;
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
 RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
 RA Kurokawa T., Nishimura O., Onda H., Fujino M.,
 RT "A prolactin-releasing peptide in the brain."
 RL Nature 393:272-276(1998).

CC -!- FUNCTION: Stimulates prolactin (PRL) release and regulates the
 expression of prolactin through its receptor GPR10. May stimulate
 lactotrophs directly to secrete PRL.
 CC -!- TISSUE SPECIFICITY: Medulla oblongata and hypothalamus.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AB015417; BAA29025.1; -;
 KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.
 FT SIGNAL 1 22
 FT PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRRP31.
 FT PROPEP 58 98 PROLACTIN-RELEASING PEPTIDE PRRP20.
 FT MOD_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).
 FT VARSPLIC 98 AA; 10544 MW; 08AC35A13B0FA908 CRC64;
 SQ SEQUENCE 98 AA; 10544 MW; 08AC35A13B0FA908 CRC64;

Query Match 95.4%; Score 103; DB 1; Length 98;
 Best Local Similarity 94.7%; Pred. No. 1.4e-09;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPDINPAWASRGIRPVGR 19
 DB 34 TPDINPAWASRGIRPVGR 52

RESULT 3

PRRP RAT STANDARD; PRT; 83 AA.

AC P81278; Q8K3Y0;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

Query Match 91.7%; Score 99; DB 1; Length 83;
 Best Local Similarity 89.5%; Pred. No. 5.2e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing
 hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-
 releasing peptide PrRP20].
 DE hormone)
 DE releasing peptide PrRP20].
 GN PRH.

OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE=Brain;
 RX MEDLINE=98268781; PubMed=9607765;
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
 RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
 RA Kurokawa T., Nishimura O., Onda H., Fujino M.,
 RT "A prolactin-releasing peptide in the brain."
 RL Nature 393:272-276(1998).

CC -!- FUNCTION: Stimulates prolactin (PRL) release and regulates the
 expression of prolactin through its receptor GPR10. May stimulate
 lactotrophs directly to secrete PRL.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=P81278-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=P81278-2; Sequence=VSP_004370;
 CC -!- TISSUE SPECIFICITY: Widely expressed, with highest levels in
 medulla oblongata and hypothalamus.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AB015418; BAA29026.1; -;
 DR EMBL; AF521930; AAM82154.1; -;
 DR PIR; JC7607; JC7607.
 KW Hormone; Amidation; Signal; Cleavage on pair of basic residues;
 FT SIGNAL 1 21 BY SIMILARITY.
 FT PEPTIDE 22 52 PROLACTIN-RELEASING PEPTIDE PRRP31.
 FT PROPEP 57 83 PROLACTIN-RELEASING PEPTIDE PRRP20.
 FT MOD_RES 52 52 AMIDATION (G-53 PROVIDE AMIDE GROUP).
 FT VARSPLIC 33 83 TPDINPAWYTGRIQVGRGRRATPRDVTGIGSCUPL
 DGRTKFSQSG -> SECLTYGKQLTSFPHFTSQMPP (in
 isoform 2).
 FT /FTId=VSP_004370.
 FT /D0C75A264EE4F29 CRC64;

Query Match 91.7%; Score 99; DB 1; Length 83;
 Best Local Similarity 89.5%; Pred. No. 5.2e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 5

ID	CYCR_CHRVI	STANDARD;	PRT;	383 AA.
AC	082947;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Photosynthetic reaction center cytochrome C subunit precursor.			
GN	PFUC.			
OS	Chromatium vinosum.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales;			
OC	Chromatiaceae; Allochromatium.			
OX	NCBI_TaxID=1049;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=D / ATCC 17899 / DSM 180;			
RA	Corson G.E., Nagashima K.V., Matsuura K., Sakuragi Y., Ruwanthi W.,			
RA	Qin H., Allen R., Knaff D.B.;			
RT	"Primary structure of genes encoding light-harvesting and reaction			
RT	center proteins from Chromatium vinosum."			
RL	Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.			
CC	-!- FUNCTION: THE REACTION CENTER OF PURPLE BACTERIA CONTAIN A			
CC	TIGHTLY BOUND CYTOCHROME MOLECULE WHICH RE-REDUCES THE PHOTO			
CC	OXIDIZED PRIMARY ELECTRON DONOR.			
CC	-!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor			
CC	(By similarity).			
CC	-!- PTM: Binds 4 heme groups per molecule.			
CC	-!- SIMILARITY: HIGH, WITH OTHER PHOTOSYNTHETIC REACTION CENTER			
CC	CYTOCHROME C SUBUNITS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; AB011811; BAA32742.1; --			
CC	HSP; P07173; GPRC.			
DR	InterPro; IPR000345; CytC_heme_BS.			
DR	InterPro; IPR003158; CytC_RC.			
DR	Pfam; PF02226; CytocRC; 1.			
DR	PIRSF; PIRSF000017; RC_cytochrome; 1.			
DR	PRODOM; PD010011; CytC RC; 1.			
DR	PROSITE; PS00190; CYTOCHROME C; 4.			
KW	Electron transport; Photosynthesis; Reaction center; Heme; Membrane;			
KW	Lipoprotein; Repeat; Signal; Palmitate.			
FT	SIGNAL 1 22 BY SIMILARITY.			
FT	CHAIN 23 383 PHOTOSYNTHETIC REACTION CENTER CYTOCHROME			
FT	LIPID 23 23 C SUBUNIT.			
FT	LIPID 23 23 S-diacylglycerol cysteine (By similarity).			
FT	METAL 94 94 IRON (HEME 1 AXIAL LIGAND)			
FT	BINDING 107 107 (BY SIMILARITY).			
FT	BINDING 110 110 HEME 1 (COVALENT) (BY SIMILARITY).			
FT	METAL 111 111 IRON (HEME 1 AXIAL LIGAND)			
FT	METAL 130 130 (BY SIMILARITY).			
FT	METAL 144 144 IRON (HEME 2 AXIAL LIGAND)			
FT	BINDING 152 152 (BY SIMILARITY).			
FT	BINDING 155 155 HEME 2 (COVALENT) (BY SIMILARITY).			
FT	METAL 156 156 IRON (HEME 2 AXIAL LIGAND)			
FT	METAL 236 236 (BY SIMILARITY).			
FT	METAL 247 247 IRON (HEME 3 AXIAL LIGAND)			
FT	BINDING 250 250 HEME 3 (COVALENT) (BY SIMILARITY).			
FT	METAL 251 251 IRON (HEME 3 AXIAL LIGAND)			

QY 1 TPDINPAWYASGIRPVGR 19

Db 33 TPDINPAWYTGIRPVGR 51

RESULT 4

ID	EX7L_STRCO	STANDARD;	PRT;	402 AA.
AC	Q9FEW3;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)			
DE	(Exonuclease VII large subunit).			
GN	XSEA OR SC05056 OR SCK7.29C.			
OS	Streptomyces coelicolor.			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Streptomycineae; Streptomycetaceae; Streptomycetes.			
OX	NCBI_TaxID=1902;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=AA3(2) / M145;			
RX	MEDLINE=21996410; PubMed=12000953;			
RA	Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,			
RA	Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,			
RA	Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,			
RA	Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,			
RA	Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,			
RA	Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,			
RA	Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,			
RA	Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,			
RA	Hopwood D.A.;			
RT	"Complete genome sequence of the model actinomycete Streptomyces			
RT	coelicolor A3(2)";			
RL	Nature 417.141-147(2002).			
CC	-!- FUNCTION: Bidirectionally degrades single-stranded DNA into large			
CC	acid-insoluble oligonucleotides, which are then degraded further			
CC	into small acid-soluble oligonucleotides (By similarity).			
CC	-!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-			
CC	or 3'- to 5'-direction to yield nucleoside 5'-phosphates.			
CC	-!- SUBUNIT: Heterooligomer composed of large and small subunits (By			
CC	similarity).			
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).			
CC	-!- SIMILARITY: Belongs to the xseA family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AL939122; CAC05901.1; --			
DR	HMAP; MF_00378; -- 1.			
DR	InterPro; IPR003753; Exonuc_VII_L.			
DR	InterPro; IPR004365; tRNA anti.			
DR	Pfam; PF02601; Exonuc_VII_L; 1.			
DR	Pfam; PF01336; tRNA anti; 1.			
DR	TIGRFAMs; TIGR00237; xseA; 1.			
KW	Hydrolase; Nuclease; Exonuclease; Complete proteome.			
SQ	SEQUENCE 402 AA; 43882 MW; 145929A8372B4E08 CRC64;			
Query Match	43.5%; Score 47; DB 1; Length 402;			
Best Local Similarity	47.6%; Pred. No. 5.1;			
Matches 10; Conservative	1; Mismatches 2; Indels 8; Gaps 1;			
QY	6 PAWYASRG-----IRPVG 18			
Db	89 PEWYAPRGQLSRAAEIKPVG 109			

FT BINDING 307 307 (BY SIMILARITY).
 FT BINDING 310 310 HEME 4 (COVALENT) (BY SIMILARITY).
 FT BINDING 311 311 HEME 4 (COVALENT) (BY SIMILARITY).
 FT METAL 311 311 IRON (HEME 4 AXIAL LIGAND)
 FT (BY SIMILARITY).
 SQ SEQUENCE 383 AA; 41522 MW; 96BCD91FF1B9AE7E CRC64;

Query Match 42.6%; Score 46; DB 1; Length 383;
 Best Local Similarity 42.1%; Pred. No. 7;
 Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 TPDINAWASRGIRPVGR 19
 ||||| : : :
 DB 263 TPQRTTAYAIRVRDQ 281

RESULT 6
 EX7L_COREF STANDARD; PRT; 413 AA.

AC Q8FQF1;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
 DE (exonuclease VII large subunit).
 GN XSEA OR CE1078.
 OS Corynebacterium efficiens.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=152794;
 RN [1]

SEQUENCE FROM N.A.
 RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
 RX MEDLINE=22723752; PubMed=12840036;
 RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
 RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
 RA Gotohori T.;
 RT Comparative complete genome sequence analysis of the amino acid
 RT replacements responsible for the thermostability of Corynebacterium
 RT efficiens.";
 RL Genome Res. 13:1572-1579 (2003).

CC -!- FUNCTION: Bidirectionally degrades single-stranded DNA into large
 CC acid-insoluble oligonucleotides, which are then degraded further
 CC into small acid-soluble oligonucleotides (By similarity).
 CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5' to 3'-
 CC or 3' to 5'-direction to yield nucleoside 5'-phosphates.
 CC -!- SUBUNIT: Heterooligomer composed of large and small subunits (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the xsea family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AP005217; BAC17898.1; ALT_INIT.
 CC HAVAP; MF_00378; -; 1.
 CC InterPro; IPR003753; Exonuc VII L.
 CC InterPro; IPR008994; Nucleic acid_OB.
 CC Pfam; PF02601; Exonuc VII L_1.
 CC TIGRFAMs; TIGR00237; xsea; 1.
 CC Hydrolase; Nuclease; Exonuclease; Complete proteome.
 SQ SEQUENCE 413 AA; 45171 MW; D3BAD9687C6BESD CRC64;

Query Match 42.6%; Score 46; DB 1; Length 413;
 Best Local Similarity 52.4%; Pred. No. 7.6;
 Matches 11; Conservative 1; Mismatches 1; Indels 8; Gaps 1;

QY 6 PAWYASRG-----IRPVG 18

DB 90 PAFVAGRGTSFLWVTDIRPVG 110
 ||:|||||
 ||:|||||

RESULT 7
 EX7L_COREF STANDARD; PRT; 417 AA.

AC Q8NRM3;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
 DE (Exonuclease VII large subunit).
 GN XSEA OR CGL1025.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1718;
 RN [1]

SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RA Nakagawa S.;
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Bidirectionally degrades single-stranded DNA into large
 CC acid-insoluble oligonucleotides, which are then degraded further
 CC into small acid-soluble oligonucleotides (By similarity).
 CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5' to 3'-
 CC or 3' to 5'-direction to yield nucleoside 5'-phosphates.
 CC -!- SUBUNIT: Heterooligomer composed of large and small subunits (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the xsea family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AP005277; BAB98418.1; -;
 CC HAVAP; MF_00378; -; 1.
 CC InterPro; IPR003753; Exonuc VII L.
 CC InterPro; IPR008994; Nucleic acid_OB.
 CC Pfam; PF02601; Exonuc VII L_1.
 CC TIGRFAMs; TIGR00237; xsea; 1.
 CC Hydrolase; Nuclease; Exonuclease; Complete proteome.
 SQ SEQUENCE 417 AA; 45582 MW; B32CD9286C173C34 CRC64;

Query Match 42.6%; Score 46; DB 1; Length 417;
 Best Local Similarity 52.4%; Pred. No. 7.6;
 Matches 11; Conservative 1; Mismatches 1; Indels 8; Gaps 1;

QY 6 PAWYASRG-----IRPVG 18
 ||:|||||
 ||:|||||

DB 94 PAFVAGRGTSFLWVTDIRPVG 114
 ||:|||||
 ||:|||||

RESULT 8
 YAAJ_ECOLI STANDARD; PRT; 476 AA.
 ID YAAJ_ECOLI
 AC P30143;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative transporter yaaJ.
 GN YAAJ OR B0007.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;


```

RN RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=92334977; PubMed=1630901;
RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
RT "Systematic sequencing of the Escherichia coli genome: analysis of
RL the 0-2.4 min region";
RN Nucleic Acids Res. 20:3305-3308(1992).
[2]
RN RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1635;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RT Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RN "The complete genome sequence of Escherichia coli K-12.";
RT Science 277:1453-1474(1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (potential).
CC -!- SIMILARITY: BELONGS TO THE SODIUM/ALANINE SYMPORTER FAMILY
CC (SAF). STRONG, TO H. INFLUENZAE HI0183.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
DR EMBL; D10483; BAB96585.1; -.
DR EMBL; AE000111; AAC73118.1; -.
DR PIR; G64720; G64720.
DR EcoGene; EGI1555; Yaaj.
DR InterPro; IPR002293; AA/rel_permease1.
DR InterPro; IPR001463; Na/Ala_symp1.
DR Pfam; PF01235; Na_Ala_symp; 1.
DR PRINTS; PR00175; NAALASMP0RT.
DR TIGRPFAMS; TIGR00835; agcs; 1.
DR PROSITE; PS00873; NA_ALANINE_SYMP; 1.
KW Hypothetical protein; Transmembrane; Inner membrane; Transport;
KW Symport; Complete proteome.
FT TRANSMEM 4 24 POTENTIAL.
FT TRANSMEM 81 101 POTENTIAL.
FT TRANSMEM 141 161 POTENTIAL.
FT TRANSMEM 174 194 POTENTIAL.
FT TRANSMEM 207 227 POTENTIAL.
FT TRANSMEM 233 253 POTENTIAL.
FT TRANSMEM 300 320 POTENTIAL.
FT TRANSMEM 351 371 POTENTIAL.
FT TRANSMEM 391 411 POTENTIAL.
FT TRANSMEM 414 434 POTENTIAL.
SQ SEQUENCE 476 AA; 51662 MW; 2F6EB2E12E126B63 CRC64;

Query Match 40.7%; Score 44; DB 1; Length 476;
Best Local Similarity 44.4%; Pred. No. 18;
Matches 8; Conservative 3; Mismatches 1; Indels 6; Gaps 1;

QY 3 DIN-----PAWVASRGI 14
|:| | | | | | | |
Db 120 DVNGQFRGGFWYMWARGL 137

RESULT 9
NRPI YEAST
ID NRPI YEAST STANDARD; PRT; 719 AA.
AC P32770; Q12228;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Asparagine-rich protein (ARP protein).

```

```

GN NRPI OR ARP1 OR ARP OR YDL167C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AH22;
RX MEDLINE=93247548; PubMed=8483449;
RA Wehner E.P., Rao E., Brendel M.;
RT "Molecular structure and genetic regulation of SFA, a gene
RT responsible for resistance to formaldehyde in Saccharomyces
RT cerevisiae, and characterization of its protein product.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Pohl T.M.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 2 RANBP2-type zinc fingers.
CC -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
DR EMBL; X68020; CAA48159.1; -.
DR EMBL; Z67750; CAA91579.1; -.
DR EMBL; Z74215; CAA98741.1; -.
DR PIR; S61046; S61046.
DR GeneOnline; 140410.
DR SGD; S0002326; NRPI.
DR InterPro; IPR000504; RNA_rec_mot.
DR InterPro; IPR001876; Znf_RandGDP.
DR Pfam; PF00076; rrm; 1.
DR Pfam; PF00641; zf-RanBP; 2.
DR SMART; SM00360; RRM; 1.
DR SMART; SM00547; Znf_RBZ; 2.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
DR PROSITE; PS01358; ZF_RANBP2_1; 2.
DR PROSITE; PS0199; ZF_RANBP2_2; 2.
KW Nuclear protein; Zinc-finger; RNA-binding; Repeat.
FT DOMAIN 226 322 RNA-BINDING (RRM).
FT ZN_FING 355 384 RANBP2-TYPE 1.
FT ZN_FING 581 610 RANBP2-TYPE 2.
FT DOMAIN 490 564 ASN-RICH.
FT CONFLICT 493 493 I -> N (IN REF. 1).
SQ SEQUENCE 719 AA; 79299 MW; ADA9BC09FD582659 CRC64;

Query Match 40.7%; Score 44; DB 1; Length 719;
Best Local Similarity 50.0%; Pred. No. 28;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 7 AWYASRGIRPVG 18
|:| | | | | |
Db 244 SWFTQYGVRPVG 255

RESULT 10
VG12_BPB03
ID VG12_BPB03 STANDARD; PRT; 860 AA.
AC Q37893;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Pre-neck appendage protein (Late protein GPI2).
GN 12.
OS Bacteriophage B103.

```



```

FT DOMAIN 1 46 EXTRACELLULAR.
FT TRANSMEM 47 71 1 (POTENTIAL).
FT DOMAIN 72 83 CYTOPLASMIC.
FT TRANSMEM 84 108 2 (POTENTIAL).
FT DOMAIN 109 123 EXTRACELLULAR.
FT TRANSMEM 124 143 3 (POTENTIAL).
FT DOMAIN 144 162 CYTOPLASMIC.
FT TRANSMEM 163 186 4 (POTENTIAL).
FT DOMAIN 187 210 EXTRACELLULAR.
FT TRANSMEM 211 238 5 (POTENTIAL).
FT DOMAIN 239 274 CYTOPLASMIC.
FT TRANSMEM 275 298 6 (POTENTIAL).
FT DOMAIN 299 306 EXTRACELLULAR.
FT TRANSMEM 307 331 7 (POTENTIAL).
FT DOMAIN 332 376 CYTOPLASMIC.
FT DISULFID 120 197 BY SIMILARITY.
FT BINDING 318 318 RETINAL CHROMOPHORE (BY SIMILARITY).
FT CARBOHYD 17 17 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 193 193 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 376 AA; 42139 MW; CCE401766AB06F26 CRC64;

Query Match 39.8%; Score 43; DB 1; Length 376;
Best Local Similarity 46.2%; Pred. No. 21;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 PDINPAWYASRGI 14
DB 40 PPMNPLWYSILGV 52

RESULT 13
OPS2_LIMPO STANDARD; PRT; 376 AA.
AC P3361;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ocellular opsin.
OS Limulus polyphemus (Atlantic horseshoe crab).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;
OC Limulidae; Limulus.
OX NCBI_TaxID=6850;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Median ocelli;
RX MEDLINE=93317641; PubMed=8327495;
RA Smith W.C., Price D.A., Greenberg R.M., Battelle B.-A.;
RT "Opsins from the lateral eyes and ocelli of the horseshoe crab,
RT Limulus polyphemus."
RL Proc. Natl. Acad. Sci. U.S.A. 90:6150-6154(1993).
CC -!- FUNCTION: Visual pigments are the light-absorbing molecules that
CC mediate vision. They consist of an apoprotein, opsin, covalently
CC linked to cis-retinal.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: OCELLAR CELLS; MEDIAN OCELLI.
CC -!- PTM: Some or all of the carboxyl-terminal Ser or Thr residues may
CC be phosphorylated (By similarity)
CC -!- MISCELLANEOUS: THIS OPSIN HAS AN ABSORPTION MAXIMUM AT 530 NM.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC Opsin subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; L03792; AAA28274.1; -
CC ENBL; L03782; AAA02499.1; -
CC PIR; A48197; A48197.
CC HSSP; P02699; 1EDV.

```

```

DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR001760; Opsin.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHOPOFSN.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL2; 1.
DR PROSITE; PS00238; OPSIN; 1.
KW Photoreceptor; Retinal protein; Transmembrane; Glycoprotein; Vision;
KW Phosphorylation; G-protein coupled receptor.
FT DOMAIN 1 46 EXTRACELLULAR.
FT TRANSMEM 47 71 1 (POTENTIAL).
FT DOMAIN 72 83 CYTOPLASMIC.
FT TRANSMEM 84 108 2 (POTENTIAL).
FT DOMAIN 109 123 EXTRACELLULAR.
FT TRANSMEM 124 143 3 (POTENTIAL).
FT DOMAIN 144 162 CYTOPLASMIC.
FT TRANSMEM 163 186 4 (POTENTIAL).
FT DOMAIN 187 210 EXTRACELLULAR.
FT TRANSMEM 211 238 5 (POTENTIAL).
FT DOMAIN 239 274 CYTOPLASMIC.
FT TRANSMEM 275 298 6 (POTENTIAL).
FT DOMAIN 299 306 EXTRACELLULAR.
FT TRANSMEM 307 331 7 (POTENTIAL).
FT DISULFID 120 197 BY SIMILARITY.
FT BINDING 318 318 RETINAL CHROMOPHORE (BY SIMILARITY).
FT CARBOHYD 17 17 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 193 193 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 376 AA; 42111 MW; PA9647C40531CBF8 CRC64;

Query Match 39.8%; Score 43; DB 1; Length 376;
Best Local Similarity 46.2%; Pred. No. 21;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 PDINPAWYASRGI 14
DB 40 PPMNPLWYSILGV 52

RESULT 14
GLGB_STRAW STANDARD; PRT; 838 AA.
AC Q82JF0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
DE enzyme) (BE) (1,4-alpha-D-glucan:1,4-alpha-D-glucan 6-glucosyl-
DE transferase).
GN GLGB OR SAV2805.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose H., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites."
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis."

```

```

RL  Nat. Biotechnol. 21:526-531(2003).
CC  -!- FUNCTION: Catalyzes the formation of the alpha-1,6-glucosidic
CC  linkages in glycogen by scission of a 1,4-alpha-linked
CC  oligosaccharide from growing alpha-1,4-glycan chains and the
CC  subsequent attachment of the oligosaccharide to the alpha-1,6
CC  position (By similarity).
CC  -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC  glycogen.
CC  -!- PATHWAY: Glycogen biosynthesis; third step.
CC  -!- SUBUNIT: Monomer (By similarity).
CC  -!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; AP005032; BAC70516.1; -.
CC  HAMAP; MF_06685; -. 1.
CC  InterPro; IPR006589; Alp_amil_cat_sub.
CC  InterPro; IPR006047; Alpha_amil_cat.
CC  InterPro; IPR006407; GLGB.
CC  Pfam; PF00128; alpha-amylase; 1_13N.
CC  Pfam; PF02922; isoamylase_N; 2.
CC  SMART; SM00842; Amy; 1.
CC  TIGRFAMs; TIGR01515; branching_enzym; 1.
CC  Glycogen biosynthesis; transferase; Glycosyltransferase;
CC  Complete proteome.
CC  ACT_SITE 408 408 BY SIMILARITY.
CC  ACT_SITE 443 443 BY SIMILARITY.
CC  ACT_SITE 448 448 BY SIMILARITY.
CC  ACT_SITE 511 511 BY SIMILARITY.
CC  ACT_SITE 513 513 BY SIMILARITY.
CC  ACT_SITE 566 566 BY SIMILARITY.
CC  ACT_SITE 633 633 BY SIMILARITY.
CC  ACT_SITE 634 634 BY SIMILARITY.
CC  ACT_SITE 838 AA; 92740 MW; 38P6F6B6AD3869CE CRC64;
CC  SEQUENCE 838 AA; 92740 MW; 38P6F6B6AD3869CE CRC64;
CC  QUERY MATCH 39.8%; Score 43; DB 1; Length 838;
CC  Best Local Similarity 72.7%; Pred. No. 46;
CC  Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
CC  -----
CC  QY 7 AWYASRGIRPV 17
CC  DB 342 AWYASRGIRPV 352
CC  -----
CC  RESULT 15
CC  ARGC_BACHD STANDARD; PRT; 345 AA.
CC  ID ARGC_BACHD STANDARD; PRT; 345 AA.
CC  AC Q9K8V2;
CC  DT 28-FEB-2003 (Rel. 41, Created)
CC  DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC  DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC  DE N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) (AGPR) (N-
CC  acetyl-glutamate semialdehyde dehydrogenase) (NAGSA dehydrogenase).
CC  GN ARGC OR BH2900.
CC  OS Bacillus halodurans.
CC  OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
CC  OX NCBI_TaxID=86665;
CC  RN [1]
CC  RP SEQUENCE FROM N.A.
CC  RC STRAIN=C-125 / JCM 9153;
CC  RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
CC  Fuji F., Hiramata C., Nakamura Y., Ogasawara N., Kuhara S.,
CC  Horikoshi K.;
CC  RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
CC  halodurans and genomic sequence comparison with Bacillus subtilis."

```

```

RL  Nucleic Acids Res. 28:4317-4331(2000).
CC  -!- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(+)
CC  + phosphate = N-acetyl-5-glutamyl phosphate + NADPH.
CC  -!- PATHWAY: Arginine biosynthesis; third step.
CC  -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC  -!- SIMILARITY: Belongs to the NAGSA dehydrogenase family. Subfamily
CC  1.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; AP001517; BAB06619.1; -.
CC  PIR; D84012; D84012.
CC  HAMAP; MF_00150; -. 1.
CC  InterPro; IPR000706; AGPR_act_site.
CC  InterPro; IPR000534; Semialdh_dh.
CC  Pfam; PF01118; Semialdh_dh; 1.
CC  Pfam; PF02774; Semialdh_dh; 1.
CC  PRODOM; PD003765; AGPR_act_site; 1.
CC  PROSITE; PS01224; ARG; 1.
CC  PROSITE; PS01224; ARG; 1.
CC  Arginine biosynthesis; Oxidoreductase; NADP; Complete proteome.
CC  ACT_SITE 149 149 BY SIMILARITY.
CC  ACT_SITE 149 149 BY SIMILARITY.
CC  SEQUENCE 345 AA; 36188 MW; 3E9P45DD09FC68EA CRC64;
CC  QUERY MATCH 39.4%; Score 42.5; DB 1; Length 345;
CC  Best Local Similarity 52.9%; Pred. No. 23;
CC  Matches 9; Conservative 1; Mismatches 6; Indels 1; Gaps 1;
CC  -----
CC  QY 2 PDINPAMYASRGIRPVG 18
CC  DB 106 PDVYEWY-KRQAPVG 121
CC  -----
CC  Search completed: August 12, 2004, 14:44:53
CC  Job time : 5.23256 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 27.5581 Seconds
(without alignments)
228.984 Million cell updates/sec

Title: US-09-700-643A-5
Perfect score: 108
Sequence: 1 TPDINPAWYASRGIRPVGRX 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL.25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_podent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103	95.4	98	Q8WN12	Q8wn12 ovis aries
2	78	72.2	117	Q9W624	Q9w624 carassius a
3	56	51.9	1120	Q7UPG7	Q7upg7 rhodospirell
4	53	49.1	750	Q9M371	Q9m371 arabidopsis
5	51	47.2	54	Q7Z6Y1	Q7z6y1 homo sapien
6	51	47.2	465	Q60687	Q60687 homo sapien
7	51	47.2	465	Q8WW85	Q8ww85 homo sapien
8	50.5	46.8	664	Q911W4	Q911w4 pseudomonas
9	50	46.3	1654	Q8FM04	Q8fm04 corynebacte
10	49	45.4	187	Q7URC6	Q7urc6 rhodospirell
11	49	45.4	402	Q821E7	Q821e7 streptomyce
12	47	43.5	284	Q50128	Q50128 pyrococcus
13	47	43.5	284	Q8U112	Q8u112 pyrococcus
14	47	43.5	299	Q87F82	Q87f82 xylella fas
15	47	43.5	333	Q9PH76	Q9ph76 xylella fas
16	47	43.5	369	Q8G4X2	Q8g4x2 bifidobacte

17	47	43.5	467	11	Q8R054	Q8r054 mus musculu
18	47	43.5	468	11	Q8K1F8	Q8k1f8 mus musculu
19	47	43.5	938	10	Q9ZVE3	Q9zve3 arabidopsis
20	46.5	43.1	333	16	Q9RJI0	Q9rji0 streptomyce
21	46.5	43.1	779	3	Q9P5J9	Q9p5j9 neurospora
22	46	42.6	145	5	Q95ZB5	Q95zb5 leishmania
23	46	42.6	419	13	Q7Z1L7	Q7z1l7 xenopus lae
24	46	42.6	419	16	Q8UKP2	Q8ukp2 agrobacteri
25	46	42.6	884	15	Q8Q7I8	Q8q7i8 human immun
26	45.5	42.1	420	2	Q93LZ7	Q93lz7 streptomyce
27	45.5	42.1	816	10	Q7XMH8	Q7xmh8 oryza sativ
28	45	41.7	179	5	Q20170	Q20170 caenorhabdi
29	45	41.7	207	5	Q27455	Q27455 brachycentr
30	45	41.7	314	16	Q7WPS0	Q7wps0 bordetella
31	45	41.7	314	16	Q7WIE3	Q7wie3 bordetella
32	45	41.7	314	16	Q7VSA1	Q7vsa1 bordetella
33	45	41.7	326	2	Q9F642	Q9f642 stigmatella
34	45	41.7	333	4	Q96SD4	Q96sd4 homo sapien
35	45	41.7	362	16	Q8DGK6	Q8dgk6 synchococc
36	45	41.7	367	4	Q8N6Q2	Q8n6q2 homo sapien
37	45	41.7	367	4	Q8IXI2	Q8ixi2 homo sapien
38	45	41.7	368	16	Q87F40	Q87f40 xylella fas
39	45	41.7	370	11	Q8CGW9	Q8cgw9 mus musculu
40	45	41.7	390	16	Q9PH18	Q9ph18 xylella fas
41	45	41.7	419	4	Q9Y276	Q9y276 homo sapien
42	45	41.7	419	4	Q7Z2V7	Q7z2v7 homo sapien
43	45	41.7	420	13	Q7ZV60	Q7zv60 brachydanlo
44	45	41.7	430	17	Q27142	Q27142 methanobact
45	45	41.7	540	10	Q9LG20	Q9lg20 oryza sativ

ALIGNMENTS

RESULT 1

Q8WN12 PRELIMINARY; PRT; 98 AA.
AC Q8WN12;
DT 01-VAR-2002 (TReMBLrel. 20, Created)
DT 01-VAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-VAR-2002 (TReMBLrel. 20, Last annotation update)
DE Preprolactin-releasing peptide.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Curlewis J.D., Kusters D.H.L., Barclay J.L., Anderson S.T.;
RT "Prolactin-releasing peptide (PrRP) in the ewe: cDNA cloning, mRNA
RT distribution and effects on prolactin secretion in vitro and in
RT vivo".
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF450453; AAL47178.1; -;
SQ SEQUENCE 98 AA; 10513 MW; 2A53331ED62CAAB5 CRC64;

Query Match 95.4%; Score 103; DB 6; Length 98;
Best Local Similarity 94.7%; Pred. No. 1e-08;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TPDINPAWYASRGIRPVGR 19
|||||
Db 34 TPDINPAWYAGRGIRPVGR 52
|||||

RESULT 2

Q9W624 PRELIMINARY; PRT; 117 AA.
ID Q9W624
AC Q9W624;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)


```

RP SEQUENCE FROM N.A.
RA Huang C.-H., Chen H., Peng J., Chen Y.;
RT "Cloning and characterization of the sushi-repeat containing protein
RT (SRP) as a novel interaction partner of Rh type C glycoprotein
RT (RhCG).";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF060567; AAC15765.1; -
DR EMBL; AF393649; AM73693.1; -
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR003410; Hyalin.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF02494; HVR; 1.
DR Pfam; PF00084; sushi; 3.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
SQ SEQUENCE 465 AA; 52971 MW; 4D752B187F3F3B8 CRC64;

Query Match 47.2%; Score 51; DB 4; Length 465;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRP 16
  |||:|||||
DB 18 TPAVTWTWAGSGYYP 33

RESULT 7
Q8W85 PRELIMINARY; PRT; 465 AA.
AC Q8W85;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Sushi-repeat protein.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020733; AAH20733.1; -
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR003410; Hyalin.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF02494; HVR; 1.
DR Pfam; PF00084; sushi; 3.
DR SMART; SM00032; CCP; 3.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
SQ SEQUENCE 465 AA; 52957 MW; 3D7229487DA1B8BD CRC64;

Query Match 47.2%; Score 51; DB 4; Length 465;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRP 16
  |||:|||||
DB 18 TPAVTWTWAGSGYYP 33

RESULT 8
Q911W4 PRELIMINARY; PRT; 664 AA.
AC Q911W4;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein PA2151.
GN PA2151.
OS Pseudomonas aeruginosa.

```

```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RE MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964 (2000).
DR EMBL; AE004642; AAG05539.1; -
DR PIR; F83376; F83376.
DR GO; GO:0003779; F:actin binding; IEA.
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR001589; Actbind_actin.
DR InterPro; IPR006047; Alpha_amyl_cat.
DR Pfam; PF00128; alpha-amylase; 1.
DR PROSITE; PS00019; ACTININ_1; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 664 AA; 76329 MW; 8F59FED54C308AD CRC64;

Query Match 46.8%; Score 50.5; DB 16; Length 664;
Best Local Similarity 62.5%; Pred. No. 20;
Matches 10; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 1 TPDINPAWYASRGIRP 16
  |||||:|
DB 478 TPDINP-WFLQSGRP 492

RESULT 9
Q8FM04 PRELIMINARY; PRT; 1654 AA.
AC Q8FM04;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Putative polyketide synthase.
GN CE2705.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA Kawarabayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
RA Usuda Y., Sugimoto S.;
RT "The entire genomic sequence of Corynebacterium efficiens YS-314.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005223; BAC19515.1; -
DR GO; GO:0016788; F:hydrolase activity, acting on ester bonds; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009058; P:biogenesis; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR InterPro; IPR001227; Ac_trans.
DR InterPro; IPR000794; Ketoacyl_synth.
DR InterPro; IPR006163; Pp_bind.
DR InterPro; IPR000379; Ser_estrse.
DR InterPro; IPR001031; Thioesterase.
DR Pfam; PF00698; Acyl_transf; 1.
DR Pfam; PF00109; ketoacyl-synt; 1.
DR Pfam; PF02801; ketoacyl-synt_C; 1.
DR Pfam; PF00550; pp-binding; 2.
DR Pfam; PF00975; Thioesterase; 1.

```

```

DR PROSITE; PS00075; ACP DOMAIN; 3.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
KW Complete proteome.
SQ SEQUENCE 1654 AA; 176466 MW; BBB60E7E3CE55CBB CRC64;

Query Match 46.3%; Score 50; DB 16; Length 1654;
Best Local Similarity 58.8%; Pred. NO. 65;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 PDINPAWYASRGIRPVG 18
DB 796 PDIEPAWYAAPGMTTVG 812
|||||:|:|:|

RESULT 10
ID Q7URC6 PRELIMINARY; PRT; 187 AA.
AC Q7URC6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN RB5749
OS Rhodospirillum rubrum.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1."
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR ENBL; BX294142; CAD7414.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 187 AA; 20737 MW; AFF81D5104D7AF21 CRC64;

Query Match 45.4%; Score 49; DB 16; Length 187;
Best Local Similarity 55.6%; Pred. NO. 8.6;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVG 18
DB 48 TPSADPMYFSAIRPHG 65
|||||:|:|:|

RESULT 11
ID Q82IE7 PRELIMINARY; PRT; 402 AA.
AC Q82IE7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative exoribonuclease large subunit.
GN XSEA OR SAV3211.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites."

```

```

Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
[2]
SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis."
RL Nat. Biotechnol. 21:526-531(2003).
DR ENBL; AP005034; BAC70322.1; -.
DR GO; GO:0009318; C:exodeoxyribonuclease VII complex; IEA.
DR GO; GO:0008855; F:exodeoxyribonuclease VII activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006308; P:DNA catabolism; IEA.
DR InterPro; IPR003753; Exonuc VII_L.
DR InterPro; IPR004365; tRNA_anti-.
DR Pfam; PF02601; Exonuc_VII_L; 1.
DR Pfam; PF01336; tRNA_anti; 1.
DR TIGRFAMs; TIGR00237; xsea; 1.
KW Complete proteome.
SQ SEQUENCE 402 AA; 43844 MW; 863AB19C4C8F4A07 CRC64;

Query Match 45.4%; Score 49; DB 16; Length 402;
Best Local Similarity 47.8%; Pred. NO. 20;
Matches 10; Conservative 1; Mismatches 2; Indels 8; Gaps 1;

QY 6 PAWYASRG-----IRPVG 18
DB 89 PENYAPRGQLSLRAAIRPIG 109
|||||:|:|:|

RESULT 12
ID O50128 PRELIMINARY; PRT; 284 AA.
AC O50128;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein PH1420.
GN PH1420.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
RL DNA Res. 5:55-76(1998).
DR ENBL; AP000006; BAA30526.1; -.
DR FR; F71015; F71015.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001330; Prenyltrans.
DR InterPro; IPR005829; Sug_Transporter.
DR InterPro; IPR008930; TerP_cyc_toroid.
DR Pfam; PF00432; prenyltrans; 2.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 284 AA; 32319 MW; 8E0E7BC3711D3815 CRC64;

Query Match 43.5%; Score 47; DB 17; Length 284;

```



```

Best Local Similarity 50.0%; Pred. No. 28;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 PDINPAWYASRGIRPVGR 19
Db 217 PYIEPTFYALRGLELGR 234

RESULT 13
Q8U112 ID Q8U112 PRELIMINARY; PRT; 284 AA.
AC Q8U112; 2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein PF1418.
GN PF1418.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb P.T., Brown J.R.;
RL "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE010244; AAL81542.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001330; P:transports.
DR InterPro; IPR005829; Sug.transports.
DR InterPro; IPR008930; Terp.cyc.toroid.
DR Pfam; PF00432; prenyltrans; 2.
DR PROSITE; PS00216; SUGAR_TRANSPORT 1; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 284 AA; 32302 MW; 1D4C5746A1931390 CRC64;

Query Match 43.5%; Score 47; DB 17; Length 284;
Best Local Similarity 50.0%; Pred. No. 28;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 PDINPAWYASRGIRPVGR 19
Db 217 PYIEPTFYALRGLELGR 234

RESULT 14
Q87F82 ID Q87F82 PRELIMINARY; PRT; 299 AA.
AC Q87F82; 2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hydroxybenzoate octaprenyltransferase.
GN UBI4 OR PD0046.
OS Xylella fastidiosa (strain Temeculal / ATCC 700964).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=183190;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=2242131; PubMed=12533478;
RA Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B.,
RA Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Moon D.H.,
RA Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva F.R.,
RA Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorry H., Tsai S.M.,
RA Carrier H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.J.,
RA Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kuramae E.E.,
RA Marino C.L., Ghiglioni E., Abreu I.L., Alves L.M.C., do Amaral A.M.,
RA Baia G.S., Blanco S.R., Brito M.S., Camnava F.S., Celestino A.V.,

```

```

da Cunha A.F., Fenille R.C., Ferro J.A., Formighieri E.F., Kishi L.T.,
RA Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Sasaki F.T., Sena J.A.D.,
RA de Souza A.A., Truffi D., Tsukumo F., Yanai G.M., Zaros L.G.,
RA Civerio E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
RA Kitajima J.P.;
RT "Comparative analyses of the complete genome sequences of Pierce's
RT disease and citrus variegated chlorosis strains of Xylella
RT fastidiosa.";
RL J. Bacteriol. 185:1018-1026(2003).
DR EMBL; AE012553; AA027953.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000537; UbiA.
DR Pfam; PF01040; UbiA; 1.
DR PROSITE; PS00943; UBI4; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 299 AA; 33965 MW; CB8C74395056D8F9 CRC64;

Query Match 43.5%; Score 47; DB 16; Length 299;
Best Local Similarity 53.3%; Pred. No. 30;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 INPAWYASRGIRPVG 18
Db 20 LDPYKWLARGDRPVG 34

RESULT 15
Q9PH76 ID Q9PH76 PRELIMINARY; PRT; 333 AA.
AC Q9PH76;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Hydroxybenzoate octaprenyltransferase.
GN XP0068.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9a5c;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.B., Kitajima J.P.,
RA Krueger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchioka M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
RL EMBL; AE003860; AAF82881.1; -.

```

DR PIR; H82852; H82852.
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0004859; F: prenyltransferase activity; IEA.
 DR GO; GO:0009058; P: biosynthesis; IEA.
 DR InterPro; IPR000537; UbiA.
 DR InterPro; IPR006370; UbiA_proteo.
 DR Pfam; PF01040; UbiA; 1.
 DR TIGRFAMs; TIGR01474; ubiA_proteo; 1.
 DR PROSITE; PS00943; UBI; 1.
 KW Complete proteome.
 SQ SEQUENCE 333 AA; 37931 MW; ECF3F4716C962B95 CRC64;
 Query Match 43.5%; Score 47; DB 16; Length 333;
 Best Local Similarity 53.3%; Pred. No. 34;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 4 INPAWYASRGIRPVG 18
 : : : : :
 Db 54 LDPYWKLGDRVVG 68

Search completed: August 12, 2004, 14:49:06
 Job time : 30.5581 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 42.4419 Seconds
(without alignments)
133.146 Million cell updates/sec

Title: US-09-700-643A-5
Perfect score: 108
Sequence: 1 TPDINPAWVARGIRPVGRX 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1980s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107	99.1	20	AAW31394	AAW31394 Human typ
2	107	99.1	20	AAW97236	AAW97236 Human typ
3	107	99.1	20	AAV49294	AAV49294 19P2 liga
4	107	99.1	20	AAB10365	AAB10365 Human oxy
5	107	99.1	20	AAB90992	AAB90992 Prolactin
6	107	99.1	20	AAG62534	AAG62534 Human CRH
7	107	99.1	20	AAE26404	AAE26404 Human PFR
8	107	99.1	20	ABU60846	ABU60846 Peptide p
9	107	99.1	21	AAW31395	AAW31395 Human typ
10	107	99.1	21	AAB10366	AAB10366 Human oxy
11	107	99.1	21	AAG62535	AAG62535 Human CRH
12	107	99.1	21	ABU60847	ABU60847 Peptide p
13	107	99.1	22	AAW31396	AAW31396 Human typ
14	107	99.1	22	AAB10367	AAB10367 Human oxy
15	107	99.1	22	AAG62536	AAG62536 Human CRH
16	107	99.1	22	ABU60848	ABU60848 Peptide p
17	107	99.1	30	AAV49299	AAV49299 19P2 liga
18	107	99.1	31	AAW31391	AAW31391 Human typ
19	107	99.1	31	AAW87615	AAW87615 Human typ
20	107	99.1	31	AAW97235	AAW97235 Human typ
21	107	99.1	31	AAV49291	AAV49291 19P2 liga
22	107	99.1	31	AAB10362	AAB10362 Human oxy
23	107	99.1	31	AAB90991	AAB90991 Prolactin
24	107	99.1	31	AAB90995	AAB90995 Prolactin
25	107	99.1	31	AAG62531	AAG62531 Human CRH

26	107	99.1	31	5	AAE26401	AAE26401 Human PFR
27	107	99.1	31	6	ABU60843	ABU60843 Peptide p
28	107	99.1	31	6	ABU60827	ABU60827 Peptide p
29	107	99.1	31	7	ADC71228	ADC71228 Human pep
30	107	99.1	32	2	AAW31392	AAW31392 Human typ
31	107	99.1	32	2	AAB10363	AAB10363 Human oxy
32	107	99.1	32	4	AAG62532	AAG62532 Human CRH
33	107	99.1	32	6	ABU60844	ABU60844 Peptide p
34	107	99.1	33	2	AAW31393	AAW31393 Human typ
35	107	99.1	33	3	AAB10364	AAB10364 Human oxy
36	107	99.1	33	4	AAG62533	AAG62533 Human CRH
37	107	99.1	33	6	ABU60845	ABU60845 Peptide p
38	107	99.1	87	2	AAW31390	AAW31390 Human typ
39	107	99.1	87	2	AAW97226	AAW97226 Human typ
40	107	99.1	87	3	AAB10361	AAB10361 Human oxy
41	107	99.1	87	4	AAG62530	AAG62530 Human CRH
42	103	95.4	19	2	AAW31370	AAW31370 Bovine G
43	103	95.4	19	2	AAW95185	AAW95185 Bovine pi
44	103	95.4	19	6	ABU60830	ABU60830 Peptide p
45	103	95.4	20	2	AAW31374	AAW31374 Bovine G

ALIGNMENTS

RESULT 1
AAW31394
ID AAW31394 standard; peptide; 20 AA.
XX

AC AAW31394;

DT 06-APR-1998 (first entry)

XX Human type G protein-coupled receptor ligand fragment 4.

XX G protein-coupled receptor; ligand binding; pharmaceutical; modulator;
KW pituitary; central nervous system; pancreas; prophylactic;
KW therapeutic agent.

XX Homo sapiens.

XX WO9724436-A2.

XX 10-JUL-1997.

XX 26-DEC-1996; 96WO-JP003821.

XX 28-DEC-1995; 95JP-00343371.

XX 15-MAR-1996; 96JP-00059419.

XX 12-AUG-1996; 96JP-00211805.

XX 18-SEP-1996; 96JP-00246573.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;

XX Kitada C;

XX WPI; 1997-363672/33.

XX N-PSDB; AAV02431.

XX Ligand peptide for G protein-coupled receptor - acts by modulating

XX function in the central nervous system, pancreas and pituitary gland.

XX Claim 2; Page 185; 258pp; English.

XX This sequence represents a peptide fragment from a novel human type

XX ligand polypeptide corresponding to amino acid residues 34 to 53 of the

XX sequence represented in AAW31390 and is used in an assay to monitor

XX ligand binding to the G protein-coupled receptor protein. Pharmaceutical

XX compositions containing this ligand may be used as a pituitary function

XX modulator, a central nervous system modulator or a pancreatic function

CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,
 CC growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,
 CC hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,
 CC ankyroplastic lateral sclerosis, acute myocardial infarction, infertility,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis and/or oligogalactia. Assays can also be developed to screen
 CC compounds which are capable of altering the binding activity of the
 CC ligand affecting activation of the G protein-coupled receptor protein
 XX
 SQ Sequence 20 AA;

Query Match 99.1%; Score 107; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.2e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPDINPAWYASRGIRPVGR 19
 |||||
 DB 1 TPDINPAWYASRGIRPVGR 19

RESULT 2
 AAW97236
 ID AAW97236 standard; peptide; 20 AA.

AC AAW97236;

XX 06-MAY-1999 (first entry)

XX Human type ligand polypeptide fragment.

XX Rat type ligand; modulation; prolactin secretion;
 KW G protein-coupled receptor; GPCR; hypocoovarianism; gonocyst cacosgenesis;
 KW menopausal syndrome; euthyroid; hypometabolism; lactation;
 KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;
 KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;
 KW contraceptive; placental function; choriocarcinoma; hydatid mole;
 KW abortion; unthrifty fetus; abnormal saccharometabolism;
 KW abnormal lipidmetabolism; oxytocia.

XX Homo sapiens.

XX WO9858962-A1.

XX 30-DEC-1998.

XX 22-JUN-1998; 98WO-JP002765.

XX 23-JUN-1997; 97JP-00165437.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Hinuma S, Kawamata Y, Fujii R, Matsumoto H;

XX WPI; 1999-105614/09.

XX Use of G protein-coupled receptor ligands - for modulating prolactin
 PT secretion or placental function, e.g. for treating menopausal syndrome,
 PT tumours, autoimmune disease or abnormal pregnancy.

XX Claim 3; Page 166; 241pp; English.

XX The present sequence represents a human type ligand fragment. It is used
 CC in the course of the invention. The specification describes an agent for
 CC modulating prolactin secretion which comprises a ligand polypeptide or a
 CC salt, for a G protein-coupled receptor (GPCR) protein. The agents for
 CC promoting prolactin secretion can be used for treating or preventing
 CC hypocoovarianism, gonocyst cacosgenesis, menopausal syndrome, euthyroid or
 CC hypometabolism. They can be used for promoting lactation in a domestic
 CC mammal and as an aphrodisiac. The agents for inhibiting prolactin

CC secretion can be used for treating or preventing pituitary adenomatosis,
 CC brain tumour, emmenopathy, autoimmune disease, prolactinoma,
 CC infertility, impotence, amenorrhea, galactorrhea, acromegaly, Chiari-
 CC Frommel syndrome, Argonz-del Castillo syndrome, Forbes-Albright syndrome,
 CC lymphoma, Sheehan syndrome or dyszoospermia. The inhibitory agents can
 CC also be used as contraceptives. The agents for modulating placental
 CC function can be used for treating or preventing choriocarcinoma, hydatid
 CC mole, irruption mole, abortion, unthrifty fetus, abnormal
 CC saccharometabolism, abnormal lipidmetabolism or oxytocia
 XX
 SQ Sequence 20 AA;

Query Match 99.1%; Score 107; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.2e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPDINPAWYASRGIRPVGR 19
 |||||
 DB 1 TPDINPAWYASRGIRPVGR 19

RESULT 3
 AAY49294
 ID AAY49294 standard; peptide; 20 AA.

AC AAY49294;

XX 22-FEB-2000 (first entry)

XX 19P2 ligand peptide fragment.

XX Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;
 KW pituitary; regulatory mechanism; central nervous system; pancreatic.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT Modified-site 20 /note= "C-terminal amide"

XX WO9960112-A1.

XX 25-NOV-1999.

XX 20-MAY-1999; 99WO-JP002650.

XX 21-MAY-1998; 98JP-00140293.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Matsumoto H, Kitada C, Hinuma S;

XX WPI; 2000-039381/03.

XX New monoclonal antibodies, useful in diagnosis, as drugs and in studying
 PT diseases related to ligand abnormality.

XX Disclosure; Page 26; 73pp; Japanese.

XX The invention provides a monoclonal antibody which has a specific
 CC reaction with the part peptide of the C-terminal of 19P2 ligand or its
 CC derivative. The antibodies can be used in diagnosis or to treat or
 CC prevent diseases associated with abnormality in the pituitary function
 CC regulatory mechanism (e.g. promotion of prolactin secretion), central
 CC nervous regulatory mechanism, and pancreatic function regulatory
 CC mechanism. The antibody-based immunoassay can also be applied in
 CC clarifying the physiological functions of the ligand and its derivative.
 CC Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand

XX Sequence 20 AA;

Query Match 99.1%; Score 107; DB 3; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.2e-10;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGR 19
 Db 1 TPDINPAWYASRGIRPVGR 19

RESULT 4
 AAB10365
 ID AAB10365 standard; peptide; 20 AA.
 XX
 AC AAB10365;
 XX
 DT 24-NOV-2000 (first entry)
 XX
 DE Human oxytocin secretion promoting peptide SEQ ID NO: 35.
 XX
 KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;
 KW veterinary medicine; milk production.
 XX
 OS Homo sapiens.
 XX
 PN WO200038704-A1.
 XX
 PD 06-JUL-2000.
 XX
 PF 22-DEC-1999; 99WO-JP007199.
 XX
 PR 25-DEC-1998; 98JP-00369585.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Matsumoto H, Kitada C, Hinuma S;
 XX
 DR WPI; 2000-452298/39.
 XX
 PT Physiologically-active polypeptide recognized as ligand by G protein-
 PT coupled receptor protein, for promoting secretion of oxytocin, as drugs
 PT for diseases relating to oxytocin secretion and in veterinary medicine.
 XX
 PS Disclosure; Page 63; 72pp; Japanese.
 XX
 CC This invention describes a novel oxytocin secretion-regulating agent
 CC which contains a ligand peptide or its salt for the G protein-coupled
 CC receptor protein. It is useful in the form of drugs for ameliorating,
 CC preventing and treating diseases relating to oxytocin secretion e.g. weak
 CC pains and atonic bleeding, before and after expulsion of placenta,
 CC uterine recovery failure, caesarean section, stoppage of artificial
 CC fertilization or galactostasis and is also applicable in veterinary
 CC medicine for promoting milk production in cow, goat and pig. This
 CC sequence represents a human peptide which acts as an oxytocin secretion
 CC promoter
 XX
 SQ Sequence 20 AA;
 Query Match 99.1%; Score 107; DB 3; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.2e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGR 19
 Db 1 TPDINPAWYASRGIRPVGR 19

RESULT 5
 AAB90992
 ID AAB90992 standard; peptide; 20 AA.
 XX
 AC AAB90992;
 XX
 DT 22-JUN-2001 (first entry)

XX DE Prolactin releasing peptide SEQ ID NO:166.
 XX
 KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
 KW blood component; modification; succinimidyl; maleimido group; amino;
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
 XX
 OS Homo sapiens.
 OS Synthetic.
 PN WO200069900-A2.
 XX
 PD 23-NOV-2000.
 XX
 PF 17-MAY-2000; 2000WO-US013576.
 XX
 PR 17-MAY-1999; 99US-0134406P.
 PR 10-SEP-1999; 99US-0153406P.
 PR 15-OCT-1999; 99US-0159783P.
 XX
 PA (CONJ-) CONJUCHEM INC.
 XX
 PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thiбаudeau K;
 XX
 DR WPI; 2001-112059/12.
 XX
 PT Modifying and attaching therapeutic peptides to albumin prevents
 PT peptidase degradation, useful for increasing length of in vivo activity.
 XX
 PS Disclosure; Page 244; 733pp; English.
 XX
 CC The present invention describes a modified therapeutic peptide (I)
 CC comprising a therapeutically active amino acid region (III) and a
 CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
 CC a less therapeutically active amino acid region (IV), which covalently
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
 CC factors and neurotransmitters, to protect them from peptidase activity in
 CC vivo for the treatment of various disorders. Endogenous therapeutic
 CC peptides are not suitable as drug candidates as they require frequent
 CC administration due to rapid degradation by peptidases in the body.
 CC Modifying and attaching therapeutic peptides to albumin prevents or
 CC reduces the action of peptidases to increase length of activity (half
 CC life) and specificity as bonding to large molecules decreases
 CC intracellular uptake and interference with physiological processes.
 CC AAS90829 to AAB92441 represent peptides which can be used in the
 CC exemplification of the present invention
 XX
 SQ Sequence 20 AA;
 Query Match 99.1%; Score 107; DB 4; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.2e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGR 19
 Db 1 TPDINPAWYASRGIRPVGR 19

RESULT 6
 AAG62534
 ID AAG62534 standard; peptide; 20 AA.
 XX
 AC AAG62534;
 XX
 DT 24-AUG-2001 (first entry)
 XX
 DE Human CRH releasing protein related peptide SEQ ID NO: 35.
 XX
 KW Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;
 KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;
 KW Addison's disease; adrenal gland hyperfunction; obesity.

XX OS Homo sapiens.
 XX PN WO200135984-A1.
 XX PD 25-MAY-2001.
 XX PF 17-NOV-2000; 2000WO-JP008119.
 XX PR 18-NOV-1999; 99JP-00327900.
 XX PR 26-SEP-2000; 2000JP-00297073.
 XX PA (TAKE) TAKEDA CHEM IND LTD.
 XX PI Kitada C, Matsumoto H, Hinuma S;
 XX DR WPI; 2001-355552/37.
 XX PT Use of G protein receptor ligand or peptide for controlling corticotropin
 PT releasing hormone secretion.
 XX FS Claim 4; Page 75; 90pp; Japanese.
 XX CC The present sequence describes a method of controlling the secretion of
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein
 CC receptor ligand. This can be used to control the secretion of CRH and is
 CC useful as an analgesic or for treating, preventing or ameliorating
 CC diseases associated with CRH secretion such as hyperaldosteronism,
 CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The
 CC present sequence is a peptide used in the exemplification of the
 CC invention
 XX SQ Sequence 20 AA;
 Query Match 99.1%; Score 107; DB 4; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.2e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TPDINPAWYASRGIRPVGR 19
 DB 1 TPDINPAWYASRGIRPVGR 19
 RESULT 7
 AA226404
 ID AAE26404 standard; peptide; 20 AA.
 XX AC AAE26404;
 XX DT 13-DEC-2002 (first entry)
 XX DE Human PrRP-31 C-terminal peptide, PrRP-20.
 XX KW Human; wakefulness; sleep disorder; prolactin releasing peptide receptor;
 XX KW PrRP; GPR10; therapy; epilepsy; narcolepsy; sleepiness; sleep apnoea;
 XX KW insomnia; idiopathic hypersomnia; psychogenic hypersomnia; seizure;
 XX KW anticonvulsant.
 XX OS Homo sapiens.
 XX PN US2002037533-A1.
 XX PD 28-MAR-2002.
 XX PF 17-AUG-2001; 2001US-00932161.
 XX PR 28-APR-2000; 2000US-00560915.
 XX PA (CIVE/) CIVELLI O.
 XX PA (LINS/) LIN S.
 XX CC

PI Civelli O, Lin S;
 XX DR WPI; 2002-403931/43.
 XX PT Screening for compounds useful for promoting wakefulness or sleep, and
 PT for treating sleeping disorders, e.g. insomnia, hypersomnia or sleep
 PT apnea, comprises administering a prolactin releasing peptide agonist or
 PT antagonist.
 XX PS Disclosure; Page 25; 35pp; English.
 XX CC The present invention relates to a method of screening for compounds for
 CC promoting wakefulness or sleep in a mammal. The method involves
 CC administering a prolactin releasing peptide (PrRP) receptor (GPR10)
 CC agonist or antagonist respectively and determining the ability of the
 CC compound to promote wakefulness or sleep. The compounds identified from
 CC the method are used in the therapy of epilepsy and other diseases
 CC associated with absence seizures and in promoting wakefulness and sleep
 CC in individuals having sleep disorders such as insomnia and narcolepsy.
 CC PrRP receptor agonists may be used to treat common disorders which lead
 CC to sleepiness, e.g. sleep apnoea, narcolepsy, idiopathic hypersomnia and
 CC psychogenic hypersomnia. PrRP receptor antagonists are useful for
 CC promoting sleep and for treating insomnia such as adjustment sleep
 CC disorder and psychophysiological insomnia. The present sequence is human
 CC PrRP-31 C-terminal peptide, PrRP-20
 XX SQ Sequence 20 AA;
 Query Match 99.1%; Score 107; DB 5; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.2e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TPDINPAWYASRGIRPVGR 19
 DB 1 TPDINPAWYASRGIRPVGR 19
 RESULT 8
 ABU60846
 ID ABU60846 standard; peptide; 20 AA.
 XX AC ABU60846;
 XX DT 06-MAY-2003 (first entry)
 XX DE Peptide production by gene recombination associated peptide #30.
 XX KW Peptide production; low-molecular peptide; KiSS-1; GPR8 ligand;
 XX KW gene recombination.
 XX OS Homo sapiens.
 XX PN WO200292829-A1.
 XX PD 21-NOV-2002.
 XX PF 16-MAY-2002; 2002WO-JP004735.
 XX PR 17-MAY-2001; 2001JP-00147341.
 XX PA (TAKE) TAKEDA CHEM IND LTD.
 XX PI Nishimura O, Suenaga M, Ito T, Kitada C;
 XX DR WPI; 2003-129302/12.
 XX PT Process for producing peptides e.g. KiSS-1 peptide and GPR8 ligand for
 PT subsequent applications by gene recombination technique through tandem
 PT repeats to provide precursor protein with specific cleavage sites.
 XX PS Disclosure; Page 68; 87pp; Japanese.
 XX CC The invention describes a method of producing a peptide comprising the

```

CC fertilization or galactostasis and is also applicable in veterinary
CC medicine for promoting milk production in cow, goat and pig. This
CC sequence represents a human peptide which acts as an oxytocin secretion
CC promoter
XX
SQ Sequence 21 AA;
      Query Match      99.1%; Score 107; DB 3; Length 21;
      Best Local Similarity 100.0%; Pred. No. 1.2e-10;
      Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TFDINPAWYASRGIRPVGR 19

```

Db 1 TPDINPAWASRGIRPVGR 19
 |||||
 RESULT 11
 AAG62535
 ID AAG62535 standard; peptide; 21 AA.
 XX
 AC AAG62535;
 XX
 DT 24-AUG-2001 (first entry)
 XX
 DE Human CRH releasing protein related peptide SEQ ID NO: 36.
 XX
 KW Human; corticotropin releasing hormone; CRH; G protein receptor ligand;
 KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;
 KW Addison's disease; adrenal gland hyperfunction; obesity.
 XX
 OS Homo sapiens.
 XX
 PN WO200135984-A1.
 XX
 PD 25-MAY-2001.
 XX
 PF 17-NOV-2000; 2000WO-JP008119.
 XX
 PR 18-NOV-1999; 99JP-00327900.
 PR 26-SEP-2000; 2000JP-00297073.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Kitada C, Matsumoto H, Hinuma S;
 XX
 DR WPI; 2001-355552/37.
 XX
 PT Use of G protein receptor ligand or peptide for controlling corticotropin
 PT releasing hormone secretion.
 XX
 PS Disclosure; Page 75; 90pp; Japanese.
 XX
 CC The present sequence describes a method of controlling the secretion of
 CC corticotropin releasing hormone (CRH), involving the use of a G protein
 CC receptor ligand. This can be used to control the secretion of CRH and is
 CC useful as an analgesic or for treating, preventing or ameliorating
 CC diseases associated with CRH secretion such as hyperaldosteronism,
 CC hypercortisolaemia, secondary or chronic hypoadrenocorticism, Addison's
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The
 CC present sequence is a peptide used in the exemplification of the
 CC invention
 XX
 SQ Sequence 21 AA;
 XX
 Query Match 99.1%; Score 107; DB 4; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.2e-10; Mismatches 0; Indels 0; Gaps 0;
 Matches 19; Conservative 0;
 QY 1 TPDINPAWASRGIRPVGR 19
 |||||
 DB 1 TPDINPAWASRGIRPVGR 19
 |||||
 RESULT 12
 ABU60847
 ID ABU60847 standard; peptide; 21 AA.
 XX
 AC ABU60847;
 XX
 DT 06-MAY-2003. (first entry)
 XX
 DE Peptide production by gene recombination associated peptide #31.
 XX
 KW Peptide production; low-molecular peptide; Kiss-1; GPR8 ligand;

KW gene recombination.
 XX
 OS Homo sapiens.
 XX
 PN WO200292829-A1.
 XX
 PD 21-NOV-2002.
 XX
 PF 16-MAY-2002; 2002WO-JP004735.
 XX
 PR 17-MAY-2001; 2001JP-00147341.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Nishimura O, Suenaga M, Ito T, Kitada C;
 XX
 DR WPI; 2003-129302/12.
 XX
 PT Process for producing peptides e.g. Kiss-1 peptide and GPR8 ligand for
 PT subsequent applications by gene recombination technique through tandem
 PT repeats to provide precursor protein with specific cleavage sites.
 XX
 PS Disclosure; Page 69; 87pp; Japanese.
 XX
 CC The invention describes a method of producing a peptide comprising the
 CC excision of the N and C-terminals of a target peptide with enzymes or
 CC chemically through the attached cleavage sites repeated by ligation in a
 CC precursor protein. The method is for producing (low-molecular) peptides
 CC e.g. Kiss-1 peptide and GPR8 ligand for subsequent applications by the
 CC gene recombination technique through tandem repeats to provide a
 CC precursor protein with specific cleavage sites. With this method, peptide
 CC production can be carried out easily to provide large quantities of the
 CC required peptides. This is the amino acid sequence of a peptide
 CC associated with the peptide production method of the invention
 XX
 SQ Sequence 21 AA;
 XX
 Query Match 99.1%; Score 107; DB 6; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.2e-10; Mismatches 0; Indels 0; Gaps 0;
 Matches 19; Conservative 0;
 QY 1 TPDINPAWASRGIRPVGR 19
 |||||
 DB 1 TPDINPAWASRGIRPVGR 19
 |||||
 RESULT 13
 AAW31396
 ID AAW31396 standard; peptide; 22 AA.
 XX
 AC AAW31396;
 XX
 DT 06-APR-1998 (first entry)
 XX
 DE Human type G protein-coupled receptor ligand fragment 6.
 XX
 KW G protein-coupled receptor; ligand binding; pharmaceutical; modulator;
 KW pituitary; central nervous system; pancreas; prophylactic;
 KW therapeutic agent.
 XX
 OS Homo sapiens.
 XX
 PN WO9724436-A2.
 XX
 PD 10-JUL-1997.
 XX
 PF 26-DEC-1996; 96WO-JP003821.
 XX
 PR 28-DEC-1995; 95JP-00343371.
 PR 15-MAR-1996; 96JP-00059419.
 PR 12-AUG-1996; 96JP-00211805.
 PR 18-SEP-1996; 96JP-00246573.
 XX

(TAKE) TAKEDA CHEM IND LTD.
 PA Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;
 PI Kitada C;
 XX
 DR N-PSDB; AAV02433.
 DR WPI; 1997-363672/33.
 XX
 XX
 PT Ligand peptide for G protein-coupled receptor - acts by modulating
 PT function in the central nervous system, pancreas and pituitary gland.
 XX
 PS Claim 2; Page 186; 258pp; English.
 XX
 XX This sequence represents a peptide fragment from a novel human type
 CC ligand polypeptide corresponding to amino acid residues 34 to 55 of the
 CC sequence represented in AAW1390 and is used in an assay to monitor
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical
 CC compositions containing this ligand may be used as a pituitary function
 CC modulator, a central nervous system modulator or a pancreatic function
 CC modulator. This ligand could have specific applications as a prophylactic
 CC or therapeutic agent for dementia, depression, hyperkinetic syndrome,
 CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,
 CC growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,
 CC hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,
 CC amyotrophic lateral sclerosis, acute myocardial infarction, infertility,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis and/or oligogalactia. Assays can also be developed to screen
 CC compounds which are capable of altering the binding activity of the
 CC ligand affecting activation of the G protein-coupled receptor protein
 XX
 XX Sequence 22 AA;
 SQ
 Query Match 99.1%; Score 107; DB 2; Length 22;
 Best Local Similarity 100.0%; Pred. No. 1.3e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TPDINPAWYASRGIRPVGR 19
 DB 1 TPDINPAWYASRGIRPVGR 19
 |||||
 RESULT 14
 AAB10367
 ID AAB10367 standard; peptide; 22 AA.
 XX
 AC AAB10367;
 XX
 DT 24-NOV-2000 (first entry)
 DE Human oxytocin secretion promoting peptide SEQ ID NO: 37.
 XX
 KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;
 KW veterinary medicine; milk production.
 XX
 OS Homo sapiens.
 XX
 PN WO200038704-A1.
 XX
 PD 06-JUL-2000.
 XX
 PF 22-DEC-1999; 99WO-JP007199.
 XX
 PR 25-DEC-1998; 98JP-00369585.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Matsumoto H, Kitada C, Hinuma S;
 XX
 DR WPI; 2000-452298/39.

XX Physiologically-active polypeptide recognized as ligand by G protein-
 PT coupled receptor protein, for promoting secretion of oxytocin, as drugs
 PT for diseases relating to oxytocin secretion and in veterinary medicine.
 XX
 PS Disclosure; Page 64; 72pp; Japanese.
 XX
 CC This invention describes a novel oxytocin secretion-regulating agent
 CC which contains a ligand peptide or its salt for the G protein-coupled
 CC receptor protein. It is useful in the form of drugs for ameliorating,
 CC preventing and treating diseases relating to oxytocin secretion e.g. weak
 CC pains and atonic bleeding, before and after expulsion of placenta,
 CC uterine recovery failure, caesarean section, stoppage of artificial
 CC fertilization or galactostasis and is also applicable in veterinary
 CC medicine for promoting milk production in cow, goat and pig. This
 CC sequence represents a human peptide which acts as an oxytocin secretion
 CC promoter
 XX
 SQ Sequence 22 AA;
 Query Match 99.1%; Score 107; DB 3; Length 22;
 Best Local Similarity 100.0%; Pred. No. 1.3e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TPDINPAWYASRGIRPVGR 19
 DB 1 TPDINPAWYASRGIRPVGR 19
 |||||
 RESULT 15
 AAG62536
 ID AAG62536 standard; peptide; 22 AA.
 XX
 AC AAG62536;
 XX
 DT 24-AUG-2001 (first entry)
 DE Human CRH releasing protein related peptide SEQ ID NO: 37.
 XX
 KW Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;
 KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;
 KW Addison's disease; adrenal gland hyperfunction; obesity.
 XX
 OS Homo sapiens.
 XX
 PN WO200135984-A1.
 XX
 PD 25-MAY-2001.
 XX
 PF 17-NOV-2000; 2000WO-JP008119.
 XX
 PR 18-NOV-1999; 95JP-00327900.
 PR 26-SEP-2000; 2000JP-00297073.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Kitada C, Matsumoto H, Hinuma S;
 XX
 DR WPI; 2001-355552/37.
 XX
 PT Use of G protein receptor ligand or peptide for controlling corticotropin
 PT releasing hormone secretion.
 XX
 PS Disclosure; Page 75; 90pp; Japanese.
 XX
 CC The present sequence describes a method of controlling the secretion of
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein
 CC receptor ligand. This can be used to control the secretion of CRH and is
 CC useful as an analgesic or for treating, preventing or ameliorating
 CC diseases associated with CRH secretion such as hyperaldosteronism,
 CC hypercortisolaemia, secondary or chronic hypoadrenocorticism, Addison's
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The

CC present sequence is a peptide used in the exemplification of the

CC invention

XX SQ Sequence 22 AA;

Query Match 99.1%; Score 107; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.3e-10;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGR 19

Db 1 TPDINPAWYASRGIRPVGR 19

Search completed: August 12, 2004, 14:43:55
Job time : 42.5669 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:49:10 ; Search time 35.3488 Seconds
(without alignments)
177.617 Million cell updates/sec

Title: US-09-700-643A-5

Perfect score: 108

Sequence: 1 TPDINPAWYASRGIRPVGRX 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1292805 segs, 313927144 residues

Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

```
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTU5_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	99.1	20	9	US-09-932-161-18
2	107	99.1	20	14	US-10-044-592-28
3	107	99.1	31	9	US-09-932-161-15
4	107	99.1	31	14	US-10-044-592-82
5	107	99.1	87	13	US-10-044-592-84
6	103	95.4	19	13	US-10-044-592-27
7	103	95.4	20	9	US-09-932-161-16
8	103	95.4	20	13	US-10-044-592-42
9	103	95.4	20	14	US-10-096-777-16
10	103	95.4	21	13	US-10-044-592-43
11	103	95.4	22	13	US-10-044-592-44
12	103	95.4	31	9	US-09-932-161-13
13	103	95.4	31	13	US-10-044-592-39
14	103	95.4	31	14	US-10-096-777-13
15	103	95.4	32	13	US-10-044-592-40

15	103	95.4	33	13	US-10-044-592-41	Sequence 41, Appl
17	103	95.4	98	13	US-10-044-592-28	Sequence 28, Appl
18	103	95.4	98	13	US-10-044-592-38	Sequence 38, Appl
19	103	95.4	98	13	US-10-044-592-82	Sequence 82, Appl
20	103	95.4	98	13	US-10-044-592-84	Sequence 84, Appl
21	103	95.4	98	13	US-10-044-592-86	Sequence 86, Appl
22	103	95.4	98	13	US-10-044-592-88	Sequence 88, Appl
23	99	91.7	20	9	US-09-932-161-17	Sequence 17, Appl
24	99	91.7	20	13	US-10-044-592-6	Sequence 6, Appl
25	99	91.7	20	14	US-10-096-777-17	Sequence 17, Appl
26	99	91.7	31	9	US-09-932-161-14	Sequence 14, Appl
27	99	91.7	31	13	US-10-044-592-4	Sequence 4, Appl
28	99	91.7	31	13	US-10-044-592-5	Sequence 5, Appl
29	99	91.7	31	14	US-10-096-777-14	Sequence 14, Appl
30	99	91.7	70	13	US-10-044-592-90	Sequence 90, Appl
31	99	91.7	82	13	US-10-044-592-1	Sequence 1, Appl
32	99	91.7	86	13	US-10-044-592-96	Sequence 96, Appl
33	99	91.7	91	13	US-10-044-592-94	Sequence 94, Appl
34	98	90.7	29	13	US-10-044-592-26	Sequence 26, Appl
35	88	81.5	25	13	US-10-044-592-78	Sequence 78, Appl
36	57	52.8	9	13	US-10-044-592-8	Sequence 8, Appl
37	51	47.2	465	14	US-10-301-822-197	Sequence 197, App
38	50.5	46.8	664	12	US-10-389-847-469	Sequence 469, App
39	49	45.4	402	14	US-10-156-761-10748	Sequence 10748, A
40	46.5	43.1	204	12	US-10-424-599-282842	Sequence 282842,
41	46	42.6	81	16	US-10-437-963-177757	Sequence 177757,
42	46	42.6	118	16	US-10-437-963-133044	Sequence 133044,
43	46	42.6	183	12	US-10-424-599-268092	Sequence 268092,
44	46	42.6	209	13	US-10-108-315-30	Sequence 30, Appl
45	46	42.6	417	9	US-09-739-626-4642	Sequence 4642, Ap

ALIGNMENTS

RESULT 1
US-09-932-161-18
; Sequence 18, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932.161
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-932-161-18

Query Match 99.1%; Score 107; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGR 19
Db 1 TPDINPAWYASRGIRPVGR 19

RESULT 2
US-10-096-777-18
; Sequence 18, Application US/10096777
; Publication No. US2003017270A1
; GENERAL INFORMATION:
; APPLICANT: Civeilli, Olivier
; APPLICANT: Lin, Steven

; TITLE OF INVENTION: Therapeutic Compositions and Methods
; FILE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)

; FILE REFERENCE: PUC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-096-777-18

Query Match 99.1%; Score 107; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGR 19
| | | | | | | | | | | | | | | | | | | | | |
Db 1 TPDINPAWYASRGIRPVGR 19

RESULT 3

US-09-932-161-15
; Sequence 15, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Civelletti, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; FILE OF INVENTION: Promoting Wakefulness and Sleep
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-932-161-15

Query Match 99.1%; Score 107; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. No. 3.1e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGR 19
| | | | | | | | | | | | | | | | | | | | | |
Db 12 TPDINPAWYASRGIRPVGR 30

RESULT 4

US-10-096-777-15
; Sequence 15, Application US/10096777
; Publication No. US20030171270A1
; GENERAL INFORMATION:
; APPLICANT: Civelletti, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; FILE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)
; FILE REFERENCE: PUC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 31

; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-096-777-15

Query Match 99.1%; Score 107; DB 14; Length 31;
Best Local Similarity 100.0%; Pred. No. 3.1e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGR 19
| | | | | | | | | | | | | | | | | | | | | |
Db 12 TPDINPAWYASRGIRPVGR 30

RESULT 5

US-10-044-592-92
; Sequence 92, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hiruma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 92
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-044-592-92

Query Match 99.1%; Score 107; DB 13; Length 87;
Best Local Similarity 100.0%; Pred. No. 8.7e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGR 19
| | | | | | | | | | | | | | | | | | | | | |
Db 34 TPDINPAWYASRGIRPVGR 52

RESULT 6

US-10-044-592-27
; Sequence 27, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hiruma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 27
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-27

Query Match

95.4%; Score 103; DB 13; Length 19;

APPLICANT: Fukusumi, Shoji
TITLE OF INVENTION: Polypeptides, their Production and Use

APPLICANT: Fukusumi, Shoji
TITLE OF INVENTION: Polypeptides, their Production and Use

Query Match 95.4%; Score 103; DB 13; Length 32;
Best Local Similarity 94.7%; Pred. No. 1.3e-08;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TPDINPAWYASRGIRPVGR 19
| | | | | | | | | | | | | | | | | | | | |
Db 12 TPDINPAWYASRGIRPVGR 30

Search completed: August 12, 2004, 15:22:51
Job time : 35.3488 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:36 ; Search time 11.3953 Seconds
(without alignments)
90.609 Million cell updates/sec

Title: US-09-700-643A-5

Perfect score: 108

Sequence: 1 TPDINPAWYASRGIRPVGR 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptcddata/2/iaa/5A COMB.pcp.*
- 2: /cgn2_6/ptcddata/2/iaa/5B COMB.pcp.*
- 3: /cgn2_6/ptcddata/2/iaa/6A COMB.pcp.*
- 4: /cgn2_6/ptcddata/2/iaa/6B COMB.pcp.*
- 5: /cgn2_6/ptcddata/2/iaa/PTUS COMB.pcp.*
- 6: /cgn2_6/ptcddata/2/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107	99.1	20	3	US-09-105-678A-46
2	107	99.1	20	3	US-08-776-971-64
3	107	99.1	20	3	US-09-421-208-46
4	107	99.1	20	4	US-09-560-915-18
5	107	99.1	21	3	US-09-105-678A-47
6	107	99.1	21	3	US-08-776-971-65
7	107	99.1	21	3	US-09-421-208-47
8	107	99.1	22	3	US-09-105-678A-48
9	107	99.1	22	3	US-08-776-971-66
10	107	99.1	22	3	US-09-421-208-48
11	107	99.1	31	3	US-09-105-678A-9
12	107	99.1	31	3	US-09-105-678A-43
13	107	99.1	31	3	US-08-776-971-61
14	107	99.1	31	3	US-09-421-208-9
15	107	99.1	31	4	US-09-421-208-43
16	107	99.1	31	4	US-09-560-915-15
17	107	99.1	32	3	US-09-105-678A-44
18	107	99.1	32	3	US-08-776-971-62
19	107	99.1	32	3	US-09-421-208-44
20	107	99.1	33	3	US-09-105-678A-45
21	107	99.1	33	3	US-08-776-971-63
22	107	99.1	33	3	US-09-421-208-45
23	107	99.1	87	3	US-08-776-971-59
24	107	99.1	87	3	US-08-776-971-135
25	107	99.1	87	3	US-08-776-971-138
26	103	95.4	19	3	US-09-105-678A-30
27	103	95.4	19	3	US-08-776-971-4

28	103	95.4	19	3	US-09-421-208-30	Sequence 30, Appl
29	103	95.4	20	3	US-09-105-678A-34	Sequence 34, Appl
30	103	95.4	20	3	US-08-776-971-8	Sequence 8, Appl
31	103	95.4	20	3	US-08-776-971-98	Sequence 98, Appl
32	103	95.4	20	3	US-09-421-208-34	Sequence 34, Appl
33	103	95.4	20	4	US-09-560-915-16	Sequence 16, Appl
34	103	95.4	21	3	US-09-105-678A-35	Sequence 35, Appl
35	103	95.4	21	3	US-08-776-971-9	Sequence 9, Appl
36	103	95.4	21	3	US-09-421-208-35	Sequence 35, Appl
37	103	95.4	22	3	US-09-105-678A-36	Sequence 36, Appl
38	103	95.4	22	3	US-08-776-971-10	Sequence 10, Appl
39	103	95.4	22	3	US-09-421-208-36	Sequence 36, Appl
40	103	95.4	31	3	US-09-105-678A-7	Sequence 7, Appl
41	103	95.4	31	3	US-09-105-678A-31	Sequence 31, Appl
42	103	95.4	31	3	US-08-776-971-5	Sequence 5, Appl
43	103	95.4	31	3	US-08-776-971-37	Sequence 37, Appl
44	103	95.4	31	3	US-09-421-208-7	Sequence 7, Appl
45	103	95.4	31	3	US-09-421-208-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1
US-09-105-678A-46
; Sequence 46, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172119/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-105-678A-46

Query Match 99.1%; Score 107; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGR 19

US-09-560-915-18

Query Match 99.1%; Score 107; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGR 19
|||||
Db 1 TPDINPAWYASRGIRPVGR 19

RESULT 5

US-09-105-678A-47
; Sequence 47, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-105-678A-47

Query Match 99.1%; Score 107; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.3e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGR 19
|||||
Db 1 TPDINPAWYASRGIRPVGR 19

RESULT 6

US-08-776-971-65
; Sequence 65, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; Habata, Yugo
; Kawamata, Yuji

Hosoya, Masaki
Fujii, Ryo
Fukusumi, Shoji
Kitada, Chieko
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,971B
FILING DATE: 06-Feb-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03821
FILING DATE: 28-DEC-1996
APPLICATION NUMBER: JP 7/343371
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: JP 8/59419
FILING DATE: 15-MAR-1996
APPLICATION NUMBER: JP 8/211805
FILING DATE: 12-AUG-1996
APPLICATION NUMBER: JP 8/246573
FILING DATE: 18-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 47176
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 65:
US-08-776-971-65

Query Match 99.1%; Score 107; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.3e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGR 19
|||||
Db 1 TPDINPAWYASRGIRPVGR 19

RESULT 7

US-09-421-208-47
; Sequence 47, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

```

; NAME: Conlin, David, G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-105-678A-48

Query Match          99.1%; Score 107; DB 3; Length 22;
Best Local Similarity 100.0%; -Pred. No. 1.4e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0;

QY      1  TPDINFAWYASRGIRPVGR 19
      |||||
Db      1  TPDINFAWYASRGIRPVGR 19

RESULT 9
US-08-776-971-66
; Sequence 66, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
;           Habata, Yugo
;           Kawamata, Yuji
;           Hosoya, Masaki
;           Fujii, Ryo
;           Fukusumi, Shoji
;           Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSER: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP.
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JF96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 66:

```

```
/
/ SEQUENCE CHARACTERISTICS:
/   LENGTH: 22 amino acids
/   TYPE: amino acid
/   STRANDEDNESS: single
/   TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ FRAGMENT TYPE: internal
/ SEQUENCE DESCRIPTION: SEQ ID NO: 66:
US-08-776-971-66

Query Match          99.1%; Score 107; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGR 19
Db 1 TPDINPAWYASRGIRPVGR 19

RESULT 10
US-09-421-208-48
; Sequence 48, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-421-208-48

Query Match          99.1%; Score 107; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGR 19
Db 1 TPDINPAWYASRGIRPVGR 19

RESULT 11
US-09-105-678A-9
; Sequence 9, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-105-678A-9

Query Match          99.1%; Score 107; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGR 19
Db 12 TPDINPAWYASRGIRPVGR 30

RESULT 12
US-09-105-678A-43
; Sequence 43, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
```

ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-43

Query Match 99.1%; Score 107; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGR 19
Db 12 TPDINPAWYASRGIRPVGR 30

RESULT 13
US-08-776-971-61
Sequence 61, Application US/08776971B
Patent No. 6228984
GENERAL INFORMATION:
APPLICANT: Hinuma, Shuji
Habata, Yugo
Kawamata, Yuji
Hosoya, Masaki
Fujii, Ryo
Fukusumi, Shoji
Kitada, Chieko
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,971B
FILING DATE: 06-Feb-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03821
FILING DATE: 28-DEC-1996
APPLICATION NUMBER: JP 7/343371
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: JP 8/59419

FILING DATE: 15-MAR-1996
APPLICATION NUMBER: JP 8/211805
FILING DATE: 12-AUG-1996
APPLICATION NUMBER: JP 8/246573
FILING DATE: 18-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 47176
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-08-776-971-61

Query Match 99.1%; Score 107; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGR 19
Db 12 TPDINPAWYASRGIRPVGR 30

RESULT 14
US-09-421-208-9
Sequence 9, Application US/09421208
Patent No. 6258561
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/421,208
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/105,678
FILING DATE: 26-JUN-1998
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids

;
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-421-208-9

Query Match 99.1%; Score 107; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGR 19
Db 12 TPDINPAWYASRGIRPVGR 30

RESULT 15

US-09-421-208-43
; Sequence 43, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208

; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48456-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440

; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-421-208-43

Query Match 99.1%; Score 107; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGR 19
Db 12 TPDINPAWYASRGIRPVGR 30

Search completed: August 12, 2004, 14:52:14
Job time : 11.3953 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: August 12, 2004, 14:37:36 ; Search time 8.83721 Seconds
(without alignments)
217.697 Million cell updates/sec

Title: US-09-700-643a-12
Perfect score: 110
Sequence: 1 TPDINPAWYAGRGIRPVGRX 20
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105	95.5	83	2 JC7607	prolactin-releasin
2	52	47.3	419	2 AH3166	hypothetical prote
3	51	46.4	430	1 B69009	conserved hypothet
4	50	45.5	527	2 T33175	hypothetical prote
5	49	44.5	790	2 T47959	hypothetical prote
6	48.5	44.1	664	2 F83376	conserved hypothet
7	46	41.8	333	2 F82852	hydroxybenzoate oc
8	46	41.8	424	2 B38176	samb protein - Sal
9	45.5	41.4	779	2 T49717	related to BCS1 pr
10	45	40.9	226	2 A87664	hypothetical prote
11	45	40.9	284	2 F71015	hypothetical prote
12	45	40.9	424	2 AB1034	UV protection prot
13	45	40.9	767	2 T21959	hypothetical prote
14	45	40.9	798	2 S11210	probable unr prote
15	44	40.0	390	2 G82844	cysteine synthase
16	44	40.0	398	2 C84780	hypothetical prote
17	43.5	39.5	506	2 H83396	probable aldehyde
18	43.5	39.5	506	2 F83142	probable aldehyde
19	43	39.1	220	2 C83292	probable glutathio
20	43	39.1	276	2 D70817	hypothetical prote
21	43	39.1	309	2 T23276	hypothetical prote
22	43	39.1	313	2 B93351	VirB6 type IV secr
23	43	39.1	359	2 C64074	fructose-bisphosph
24	43	39.1	399	2 T30222	sensory protein ki
25	43	39.1	476	2 G64720	probable amino aci
26	43	39.1	476	2 G90629	probable inner mem
27	43	39.1	476	2 G85480	inner membrane tra
28	43	39.1	503	2 A82193	Sun/nucleolar prot
29	43	39.1	511	2 B37222	cytochrome P450 1A

30	43	39.1	548	2 T47548	hypothetical prote
31	43	39.1	627	2 H96951	fusion, PTS system
32	43	39.1	938	2 C84480	hypothetical prote
33	43	39.1	1292	2 T31462	probable magnesium
34	42	38.2	143	2 T36978	probable transposa
35	42	38.2	284	2 A75117	hypothetical prote
36	42	38.2	311	2 S66600	cytochrome-c oxida
37	42	38.2	335	2 B87551	lytB homolog limpo
38	42	38.2	352	2 B87507	conserved hypothet
39	42	38.2	375	2 F91173	probable transport
40	42	38.2	375	2 F86019	probable transport
41	42	38.2	375	2 S47704	hypothetical 41.1K
42	42	38.2	428	2 F81694	pyruvate dehydroge
43	42	38.2	433	2 H87660	peptidoglycan-bind
44	42	38.2	453	2 S18997	tubulin beta chain
45	42	38.2	468	2 C83160	nitrite extrusion

ALIGNMENTS

RESULT 1

JC7607
prolactin-releasing peptide - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C:Accession: JC7607
R:Amada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Hosoya, T
Biochem. Biophys. Res. Commun. 281, 53-56, 2001
A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene: Mu
A:Reference number: JC7607; MUID:21092785; PMID:11178959
A:Contents: Spleen
A:Accession: JC7607
A:Molecule type: DNA
A:Residues: 1-83 <YAM>
A:Cross-references: DBJ:AB040612; DBJ:AB040613
C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior p
release, and stimulation of ACTH secretion from the pituitary.
C:Genetics:
A:Gene: PrRP
A:Introns: 33/1

Query Match 95.5%; Score 105; DB 2; Length 83;
Best Local Similarity 94.7%; Pred. No. 4.9e-09;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY	1	TPDINPAWYAGRGIRPVGR	19
DB	33	TPDINPAWYTGGRPVGR	51

RESULT 2

AH3166
hypothetical protein ugpB [imported] - Agrobacterium tumefaciens (strain C58, Dupont) p
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C:Accession: AH3166
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, Y
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AH3166
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-419 <KUR>
A:Cross-references: GB:AE008697; PIDN:RAL45750.1; PID:gi7743483; GSPDB:GN00188
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: ugpB

A:Cross-references: GB:AE003860; GB:AE003849; NID:g9104930; PIDN:AAF2881.1; GSPDB:GN001
A:Experimental source: strain 95c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorzy, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krueger, J.E.; Kuranae, E.E.; Laigh
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0068
C:Superfamily: 4-hydroxybenzoate octaprenyltransferase
Query Match 41.8%; Score 46; DB 2; Length 333;
Best Local Similarity 53.3%; Pred. No. 21;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 4 INPAWYAGRGIRPVG 18
DB 54 LDPYKLGADGDRPVG 68
RESULT 8
Samb protein - Salmonella typhimurium
C:Species: Salmonella typhimurium
C:Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 20-Jun-2000
C:Accession: B38176
R:Nohmi, T.; Nakai, Y.; Watanabe, M.; Murayama, S.Y.; Sofuni, T.
J. Bacteriol. 173, 1051-1063, 1991
A:Title: Salmonella typhimurium has two homologous but different umuDC operons: cloning
A:Reference number: A38176; MUID:91123176; PMID:1991707
A:Accession: B38176
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-424 <NOH>
A:Cross-references: GB:D90202; NID:g217097; PIDN:BAAL4226.1; PID:g217089
A:Experimental source: strain LT2
C:Genetics:
A:Gene: samb
C:Function:
A:Description: restores UV mutability; involved in mutagenesis
C:Superfamily: umuC protein
C:Keywords: DNA repair; induced mutagenesis; SOS mutagenesis
Query Match 41.8%; Score 46; DB 2; Length 424;
Best Local Similarity 53.3%; Pred. No. 27;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
QY 2 PDINPAWYAGRGIRP 16
DB 384 PGKGKINPAGRGIRP 398
RESULT 9
related to BCS1 protein precursor [imported] - Neurospora crassa
N:Alternate names: protein B23121.300
C:Species: Neurospora crassa
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
A:Accession: T49717
R:Schulte, U.; Aigt, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25022
A:Accession: T49717
A>Status: preliminary

A:Molecule type: DNA
A:Residues: 1-779 <SCH>
A:Cross-references: EMBL:AL356172; GSPDB:GN00116; NCSP:B23L21.300
A:Experimental source: BAC clone B23L21; strain OR74A
C:Genetics:
A:Gene: NCSP:B23L21.300
A:Map position: 6
Query Match 41.4%; Score 45.5; DB 2; Length 779;
Best Local Similarity 58.8%; Pred. No. 59;
Matches 10; Conservative 1; Mismatches 3; Indels 3; Gaps 1;
QY 1 TPDINPA---WYAGRGI 14
DB 286 TDYLNPAATRRRYANRGI 302
RESULT 10
A87664
hypothetical protein CC3347 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: A87664
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolor
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: A87664
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-226 <STO>
A:Cross-references: GB:AE005673; NID:g13425049; PIDN:AAK25309.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC3347
Query Match 40.9%; Score 45; DB 2; Length 226;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 4 INPAWYAGRGIRPV 17
DB 125 VNPDRSGRALRDV 138
RESULT 11
F71015
hypothetical protein PH1420 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000
C:Accession: F71015
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekir
M.; Chikuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kishida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: F71015
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-284 <KAW>
A:Cross-references: GB:AP000006; NID:g3236133; PIDN:BAA30526.1; PID:g3257843
A:Experimental source: strain OT3
A>Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH1420
C:Superfamily: Pyrococcus horikoshii hypothetical protein PH1420
Query Match 40.9%; Score 45; DB 2; Length 284;
Best Local Similarity 50.0%; Pred. No. 26;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
QY 2 PDINPAWYAGRGIRPVGR 19

Db 217 PYIEPTFVALRGLLEGR 234
| | | | | | | | | |

RESULT 12

A:Title: UV protection protein [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain 14024)
A:Reference number: S11210; MUID:90370473; PMID:2204029
A:Accession: S11210
A:Molecule type: mRNA
A:Residues: 1-798 <JEFF>
A:Cross-references: EMBL:X52311; NID:g57454; PIDN:CAA36549.1; PID:g57455
C:Keywords: DNA binding

Query Match 40.9%; Score 45; DB 2; Length 798;
Best Local Similarity 47.1%; Pred. No. 73;
Matches 8; Conservative 5; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPV 17
| | | | | | | | | |
Db 589 TEANPTIYSGKVIREF 605
| | | | | | | | | |

RESULT 15

A:Title: cysteine synthase XF0128 [imported] - Xylella fastidiosa (strain 9a5c)
A:Reference number: S11210; MUID:90370473; PMID:2204029
A:Accession: S11210
A:Molecule type: mRNA
A:Residues: 1-798 <JEFF>
A:Cross-references: EMBL:X52311; NID:g57454; PIDN:CAA36549.1; PID:g57455
C:Keywords: DNA binding

Query Match 40.9%; Score 45; DB 2; Length 798;
Best Local Similarity 47.1%; Pred. No. 73;
Matches 8; Conservative 5; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPV 17
| | | | | | | | | |
Db 589 TEANPTIYSGKVIREF 605
| | | | | | | | | |

Query Match 40.9%; Score 45; DB 2; Length 424;
Best Local Similarity 77.8%; Pred. No. 38;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 WYAGRGIRP 16
| | | | | | | | | |

Db 390 WFAGRGIRP 398
| | | | | | | | | |

RESULT 13

A:Title: hypochlorite protein F38E11.7 - Caenorhabditis elegans
A:Reference number: S11210; MUID:90370473; PMID:2204029
A:Accession: S11210
A:Molecule type: DNA
A:Residues: 1-767 <WIL>
A:Cross-references: EMBL:Z68342; PIDN:CAA92775.1; GSPDB:GN00022; CESP:F38E11.7
A:Experimental source: Clone F38E11
A:Genetics: T21969
A:Map position: 4
A:Introns: 50/2; 118/1; 139/2; 169/3; 226/1; 248/1; 287/2; 375/2; 432/3; 465/3; 548/1; 6

Query Match 40.9%; Score 45; DB 2; Length 767;
Best Local Similarity 75.0%; Pred. No. 70;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PDINPAWY 9
| | | | | | | | | |

Db 747 PDVPAWY 754
| | | | | | | | | |

RESULT 14

A:Title: probable untr protein - rat
A:Reference number: S11210; MUID:90370473; PMID:2204029
A:Accession: S11210
A:Molecule type: DNA
A:Residues: 1-767 <WIL>
A:Cross-references: EMBL:Z68342; PIDN:CAA92775.1; GSPDB:GN00022; CESP:F38E11.7
A:Experimental source: Clone F38E11
A:Genetics: T21969
A:Map position: 4
A:Introns: 50/2; 118/1; 139/2; 169/3; 226/1; 248/1; 287/2; 375/2; 432/3; 465/3; 548/1; 6

Query Match 40.9%; Score 45; DB 2; Length 767;
Best Local Similarity 75.0%; Pred. No. 70;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PDINPAWY 9
| | | | | | | | | |
Db 747 PDVPAWY 754
| | | | | | | | | |

Search completed: August 12, 2004, 14:50:29
Job time : 9.83721 secs

Query Match 40.0%; Score 44; DB 2; Length 390;
Best Local Similarity 70.0%; Pred. No. 50;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 NPAWYAGRG 14
| | | | | | | | | |

Db 354 SPAWYAAHGI 363
| | | | | | | | | |

Result No.	Score	Query Match	Length	DB	ID	Description
1	109	99.1	98	1	PRRP_BOVIN	P81264 bos taurus
2	105	95.5	87	1	PRRP_RAT	P81278 rattus norv
3	103	93.6	83	1	PRRP_HUMAN	P81277 homo sapien
4	52	47.3	413	1	EXYL_COREF	Q8fcp1 corynebacte
5	52	47.3	417	1	EXYL_CORGL	Q8nrm3 corynebacte
6	46	41.8	402	1	EXYL_STRCO	Q9fbm3 streptomyc
7	46	41.8	424	1	SAMB_SALTY	P23832 salmonella
8	45	40.9	798	1	UNR_RAT	P81395 rattus norv
9	44	40.0	383	1	CYCR_CHRVI	O82947 chromatium
10	43	39.1	359	1	ALF_HABIN	P44429 haemophilus
11	43	39.1	476	1	YAAV_ECOLI	P30143 escherichia
12	43	39.1	511	1	CP12_CANFA	P55592 canis fami
13	42	38.2	272	1	TRA2_DROVI	O20008 drosophila
14	42	38.2	332	1	ISPH_MTCLE	Q9x781 mycobacteri
15	42	38.2	374	1	YHHU_ECOLI	P31993 escherichia
16	42	38.2	453	1	TBB2_GSOON	P32925 geotrichum
17	42	38.2	497	1	AMPA_RHIME	Q923y7 rhizobium m
18	42	38.2	719	1	NRPI_YEAST	P32770 saccharomy
19	42	38.2	938	1	GLGE_STRAW	Q821f0 streptomyc
20	41.5	37.7	345	1	ARGC_BACHD	Q9K8v2 bacillus ba
21	41.5	37.7	652	1	TETE_CLOPE	Q46306 clostridium
22	41	37.3	211	1	GIDB_PBSHM	Q8rct4 pseudomonas
23	41	37.3	302	1	HMEB_RALSO	Q8xxc3 ralstonia s
24	41	37.3	313	1	SPEB_STRCL	P37819 streptomyc
25	41	37.3	342	1	Y762_METJA	Q5a172 methanococc
26	41	37.3	347	1	Y576_METJA	Q57996 methanococ
27	41	37.3	359	1	ALRI_YERPE	Q8x110 yersinia pe
28	41	37.3	376	1	OPSI_LIMPO	P35360 limulus pol
29	41	37.3	376	1	OPSI_LIMPO	P35361 limulus pol
30	41	37.3	391	1	GATS_CHICK	P43692 gallus gall
31	41	37.3	428	1	NER3_HUMAN	Q9uq49 homo sapien
32	41	37.3	462	1	ATR2_HUMAN	P58335 homo sapien
33	41	37.3	569	1	ATR1_MOUSE	Q9cz52 mus musculi

```

RESULT 2
PRRP RAT
ID PRRP RAT STANDARD; PRT; 83 AA.
AC P81278; Q8K370;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing
DE hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-
DE releasing peptide PrRP20].
GN PRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=96268781; PubMed=9607765;
RA Hiruma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
RT "A prolactin-releasing peptide in the brain.";
RL Nature 393:272-276(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=Sprague-Dawley; TISSUE=Hypothalamus;
RA Anderson S.T., Kokay I.C., Lang T., Grattan D.R., Culewis J.D.;
RT "Quantitation of prolactin-releasing peptide (PrRP) mRNA expression in
RT specific brain regions during the rat oestrous cycle and in
RT lactation.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP TISSUE SPECIFICITY.
RX MEDLINE=99426652; PubMed=10498338;
RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,
RA Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,
RA Sunino Y., Fujino M.;
RT "Tissue distribution of prolactin-releasing peptide (PrRP) and its
RT receptor.";
RL Regul. Pept. 83:1-10(1999).
CC -!- FUNCTION: Stimulates prolactin (PRL) release and regulates the
CC expression of prolactin through its receptor GPR10. May stimulate
CC lactotrophs directly to secrete PRL.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P81278-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P81278-2; Sequence=VSP 004370;
CC -!- TISSUE SPECIFICITY: Widely expressed, with highest levels in
CC medulla oblongata and hypothalamus.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AB015418; BAA29026.1; -.
CC DR EMBL; AF521930; AAM82154.1; -.
CC DR PIR; JC7607; JC7607.
CC KW Hormone; Amidation; Signal; Cleavage on pair of basic residues;
CC Alternative splicing.
CC FT SIGNAL 1 21 BY SIMILARITY.
CC FT PEPTIDE 22 52 PROLACTIN-RELEASING PEPTIDE PRRP31.
CC FT PROPEP 33 52 PROLACTIN-RELEASING PEPTIDE PRRP20.
CC FT MOD RES 57 83
CC FT VARSPLIC 52 83
CC FT AMIDATION (G-53 PROVIDE AMIDE GROUP).
CC TPDINPAWYTGIRPVGRRATPRDVTGLQSLCLPL
CC DORTXFSQSG -> SECLTYGKQPLTFSPFTSQMPP (in

```

```

isoform 2).
FT FT /FTid=VSP 004370.
SQ SQ DQC75A264EEE4F29 CRC64;
Query Match 95.5%; Score 105; DB 1; Length 83;
Best Local Similarity 94.7%; Pred. No. 1.5e-09;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TPDINPAWYTGIRPVGRR 19
| | | | | | | | | | | | | | | | | |
DB 33 TPDINPAWYTGIRPVGRR 51
| | | | | | | | | | | | | | | | | |
RESULT 3
PRRP HUMAN
ID PRRP HUMAN STANDARD; PRT; 87 AA.
AC P81277;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing
DE hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-
DE releasing peptide PrRP20].
GN PRH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98268781; PubMed=9607765;
RA Hiruma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
RT "A prolactin-releasing peptide in the brain.";
RL Nature 393:272-276(1998).
RN [2]
RP TISSUE SPECIFICITY.
RX MEDLINE=99426652; PubMed=10498338;
RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,
RA Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,
RA Sunino Y., Fujino M.;
RT "Tissue distribution of prolactin-releasing peptide (PrRP) and its
RT receptor.";
RL Regul. Pept. 83:1-10(1999).
CC -!- FUNCTION: Stimulates prolactin (PRL) release and regulates the
CC expression of prolactin through its receptor GPR10. May stimulate
CC lactotrophs directly to secrete PRL.
CC -!- TISSUE SPECIFICITY: Medulla oblongata and hypothalamus.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AB015419; BAA29027.1; -.
CC DR EMBL; AB015418; BAA29026.1; -.
CC DR GO; GO:0005180; F:peptide hormone; TAS.
CC KW Hormone; Amidation; Signal.
CC FT SIGNAL 1 22 BY SIMILARITY.
CC FT PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRRP31.
CC FT PROPEP 34 53 PROLACTIN-RELEASING PEPTIDE PRRP20.
CC FT MOD RES 58 87
CC FT AMIDATION (G-54 PROVIDE AMIDE GROUP).
CC TPDINPAWYTGIRPVGRRATPRDVTGLQSLCLPL
CC DORTXFSQSG -> SECLTYGKQPLTFSPFTSQMPP (in

```

```

Query Match 93.6%; Score 103; DB 1; Length 87;
Best Local Similarity 94.7%; Pred. No. 3.2e-09;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```
QY 1 TPDINPAWYAGRGIRPVGR 19
DE |||||||
DB 34 TPDINPAWYASRGIRPVGR 52
|||
RESULT 4
EX7L COREF STANDARD; PRT; 413 AA.
ID -EX7L COREF STANDARD; PRT; 413 AA.
AC Q8FOPI;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
DE (Exonuclease VII large subunit).
GN XSEA OR CGL1025.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RX MEDLINE=22723752; PubMed=12840036;
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens."
RL Genome Res. 13:1572-1579 (2003).
CC -!- FUNCTION: Bidirectionally degrades single-stranded DNA into large
CC acid-insoluble oligonucleotides, which are then degraded further
CC into small acid-soluble oligonucleotides (By similarity).
CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.
CC -!- SUBUNIT: Heterooligomer composed of large and small subunits (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the xsea family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AP005217; BAB17888.1; ALT_INIT.
CC HAVAP; MF 00378; 1.
CC InterPro: IPR003753; Exonuc_VII_L.
CC InterPro: IPR008994; Nucleic_acid_OB.
CC Pfam: PF02601; Exonuc_VII_L; 1.
CC TIGRFAMs: TIGR00237; xsea; 1.
CC Hydrolase; Nuclease; Exonuclease; Complete proteome.
CC SEQUENCE 413 AA; 45171 MW; D3BADD9687C6EED CRC64;
CC
Query Match 47.3%; Score 52; DB 1; Length 413;
Best Local Similarity 57.1%; Pred. No. 1.2;
Matches 12; Conservative 1; Mismatches 0; Indels 8; Gaps 1;
QY 6 PAWYAGRG-----IRPVG 18
DB 90 PAFYAGRGPSLWYTDIRPVG 110
|||||
RESULT 5
EX7L CORGL STANDARD; PRT; 417 AA.
ID -EX7L CORGL STANDARD; PRT; 417 AA.
AC Q8RNM3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
DE (Exonuclease VII large subunit).
GN XSEA OR SC05056 OR SCK7.29C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA
```

RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 CC -!- FUNCTION: Bidirectionally degrades single-stranded DNA into large
 CC acid-insoluble oligonucleotides, which are then degraded further
 CC into small acid-soluble oligonucleotides (By similarity).
 CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-
 CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.
 CC -!- SUBUNIT: Heterooligomer composed of large and small subunits (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the xseA family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AL939122; CAC05901.1; -.
 CC HAMAP; MF 00378; -; 1.
 CC InterPro; IPR003753; Exonuc VII_L.
 CC InterPro; IPR004365; tRNA anti.
 CC Pfam; PF02601; Exonuc_VII_L; 1.
 CC Pfam; PF01336; tRNA anti; 1.
 CC TIGRFAMs; TIGR00237; xseA; 1.
 CC HydroLase; Nuclease; Exonuclease; Complete proteome.
 SQ SEQUENCE 402 AA; 43882 MW; 145929A6372B4E08 CRC64;
 CC
 CC Query Match 41.8%; Score 46; DB 1; Length 402;
 CC Best Local Similarity 47.6%; Pred. No. 9.8;
 CC Matches 10; Conservative 1; Mismatches 2; Indels 8; Gaps 1;
 CC
 CC 6 PANWAGRG-----IRPVG 18
 CC |||||
 CC 89 PENWARGQLRAAEIKPVG 109
 CC |||||
 CC
 CC RESULT 7
 CC SAMB SALTY STANDARD; PRT; 424 AA.
 CC AC P23832;
 CC DT 01-NOV-1991 (Rel. 20, Created)
 CC DT 01-NOV-1991 (Rel. 20, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE SAMB protein.
 CC GN SAMB.
 CC OS Salmonella typhimurium.
 CC OG Plasmid 60-mDa cryptic.
 CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC OC Enterobacteriaceae; Salmonella.
 CC OX NCBI_TaxID=602;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC RP STRAIN=LT2;
 CC RX MEDLINE=91123176; PubMed=191707;
 CC RA Nohmi T., Hakura A., Nakai Y., Watanabe M., Murayama S.Y.,
 CC RA Sofuni T.;
 CC RT "Salmonella typhimurium has two homologous but different umuDC
 CC RT operons: cloning of a new umuDC-like operon (samAB) present in a
 CC RT 60-megadalton cryptic plasmid of S. typhimurium.";
 CC RL J. Bacteriol. 173:1051-1063(1991).
 CC CC -!- FUNCTION: Involved in UV protection and mutation.
 CC -!- SIMILARITY: Belongs to the DNA polymerase type-Y family.
 CC -!- SIMILARITY: Contains 1 umuC domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D90202; BAAL4226.1; -.
 CC PIR; B38176; B38176.
 CC HAMAP; MF 01113; atypical; 1.
 CC InterPro; IPR001126; UMU_C_like.
 CC Pfam; PF00817; IMS; 1.
 CC PROSITE; PS00173; UMU_C; 1.
 CC PROSITE; PS00352; COLD_SHOCK; 4.
 CC PROSITE; PS00357; CSP; 5.
 CC RNA-binding; Repeat. 87
 CC DOMAIN 26 CSD 1.
 CC FT DOMAIN 136 179 CSD 2 (INCOMPLETE).
 CC FT DOMAIN 186 245 CSD 3.
 CC FT DOMAIN 297 337 CSD 4 (INCOMPLETE).
 CC FT DOMAIN 349 410 CSD 5.
 CC
 CC RESULT 8
 CC UNR_RAT STANDARD; PRT; 798 AA.
 CC ID UNR_RAT
 CC AC P18395;
 CC DT 01-NOV-1990 (Rel. 16, Created)
 CC DT 01-NOV-1990 (Rel. 16, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE UNR protein.
 CC GN UNR.
 CC OS Rattus norvegicus (Rat).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC OX NCBI_TaxID=10116;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC RP TISSUE=Testis;
 CC RX MEDLINE=90370473; PubMed=2204029;
 CC RA Jeffers M., Paciucci R., Pellicer A.;
 CC RT "Characterization of unr, a gene closely linked to N-ras";
 CC RL Nucleic Acids Res. 18:4891-4899(1990).
 CC CC -!- FUNCTION: RNA-binding protein (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to the cold-shock domain (CSD) family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X52311; CAA36549.1; -.
 CC PIR; S11210; S11210.
 CC HSP; P15277; IMJC.
 CC InterPro; IPR002059; Cold_shock.
 CC InterPro; IPR008994; Nucleic_acid_OB.
 CC Pfam; PF00313; CSD; 7.
 CC ProDom; PD000621; Cold_shock; 1.
 CC SMART; SM00357; CSP; 5.
 CC PROSITE; PS00352; COLD_SHOCK; 4.
 CC RNA-binding; Repeat. 87
 CC DOMAIN 26 CSD 1.
 CC FT DOMAIN 136 179 CSD 2 (INCOMPLETE).
 CC FT DOMAIN 186 245 CSD 3.
 CC FT DOMAIN 297 337 CSD 4 (INCOMPLETE).
 CC FT DOMAIN 349 410 CSD 5.


```

FT DOMAIN 447 507 CSD 6.
FT DOMAIN 519 579 CSD 7.
FT DOMAIN 610 670 CSD 8.
FT DOMAIN 674 735 CSD 9.
SQ SEQUENCE 798 AA; 88894 MW; F484B3FA8B0995A4 CRC64;

Query Match 40.9%; Score 45; DB 1; Length 798;
Best Local Similarity 47.1%; Pred. No. 26;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPV 17
   ||| ||| |||
Db 589 TEANPTVYSGKVRPL 605

RESULT 9
CYCR_CHRVI
ID CYCR_CHRVI STANDARD; PRT; 383 AA.
AC O82947;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Photosynthetic reaction center cytochrome C subunit precursor.
GN PUF.
OS Chromatium vinosum.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales;
OC Chromatiaceae; Allochromatium.
OX NCBI_TaxID=1049;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=D / ATCC 17899 / DSM 180;
RA Corson G.E., Nagashima K.V., Matsuura K., Sakuragi Y., Ruwanchi W.,
RA Qin H., Allen R., Knaff D.B.;
RT "Primary structure of genes encoding light-harvesting and reaction
RT center proteins from Chromatium vinosum."
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THE REACTION CENTER OF PURPLE BACTERIA CONTAIN A
CC TIGHTLY BOUND CYTOCHROME MOLECULE WHICH RE-REDUCES THE PHOTO
CC OXIDIZED PRIMARY ELECTRON DONOR.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (By similarity).
CC -1- PTM: Binds 4 heme groups per molecule.
CC -1- SIMILARITY: HIGH, WITH OTHER PHOTOSYNTHETIC REACTION CENTER
CC CYTOCHROME C SUBUNITS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AB011811; BAA32742.1; -.
CC DR HSPB; P07173; 6PRC
CC DR InterPro; IPR000345; CytC_heme_BS.
CC DR InterPro; IPR003158; CytC_RC.
CC DR Pfam; PF02276; Cytoc RC; 1.
CC DR PROSITE; PROSITE00017; RC cytochrome; 1.
CC DR PROSITE; PS00101; CytC RC; 1.
CC DR PROSITE; PS00190; CYTOCHROME C; 4.
CC DR Electron transport; Photosynthesis; Reaction center; Heme; Membrane;
CC Lipoprotein; Repeat; Signal; Palmitate
CC SIGNAL 1 22
CC FT CHAIN 23 383
CC PHOTOSYNTHETIC REACTION CENTER CYTOCHROME
CC C SUBUNIT.
CC S-diacylglycerol cysteine (By
CC similarity).
CC LIPID 23 23
CC LIPID 23 23 N-palmitoyl cysteine (By similarity).
CC FT METAL 94 94
CC IRON (HEME 1 AXIAL LIGAND)
CC (BY SIMILARITY).
CC FT BINDING 107 107
CC HEME 1 (COVALENT) (BY SIMILARITY).
CC FT BINDING 110 110
CC HEME 1 (COVALENT) (BY SIMILARITY).

```

```

FT METAL 111 111
   IRON (HEME 1 AXIAL LIGAND)
   (BY SIMILARITY).
FT METAL 130 130
   IRON (HEME 2 AXIAL LIGAND)
   (BY SIMILARITY).
FT METAL 144 144
   IRON (HEME 4 AXIAL LIGAND)
   (BY SIMILARITY).
FT BINDING 152 152
   HEME 2 (COVALENT) (BY SIMILARITY).
FT BINDING 155 155
   HEME 2 (COVALENT) (BY SIMILARITY).
FT METAL 156 156
   IRON (HEME 2 AXIAL LIGAND)
   (BY SIMILARITY).
FT METAL 236 236
   IRON (HEME 3 AXIAL LIGAND)
   (BY SIMILARITY).
FT BINDING 247 247
   HEME 3 (COVALENT) (BY SIMILARITY).
FT BINDING 250 250
   IRON (HEME 3 AXIAL LIGAND)
   (BY SIMILARITY).
FT METAL 251 251
   IRON (HEME 3 AXIAL LIGAND)
   (BY SIMILARITY).
FT BINDING 307 307
   HEME 4 (COVALENT) (BY SIMILARITY).
FT BINDING 310 310
   HEME 4 (COVALENT) (BY SIMILARITY).
FT METAL 311 311
   IRON (HEME 4 AXIAL LIGAND)
   (BY SIMILARITY).
SQ SEQUENCE 383 AA; 41522 MW; 96BCD91FF1B9AE7E CRC64;

Query Match 40.0%; Score 44; DB 1; Length 383;
Best Local Similarity 42.1%; Pred. No. 19;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
   ||| ||| ||| |||
Db 263 TPQRTTAWYAIRHVRDINQ 281

RESULT 10
ALF_HAEIN
ID ALF_HAEIN STANDARD; PRT; 359 AA.
AC P44429;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Fructose-bisphosphate aldolase (EC 4.1.2.13).
GN FBA OR H10524.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.B., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
RA rd."
RL Science 269:496-512(1995).
CC -1- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glycerone,
CC phosphate + D-glyceraldehyde 3-phosphate.
CC -1- COFACTOR: Zinc (By similarity).
CC -1- PATHWAY: Glycolysis; sixth step.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SIMILARITY: Belongs to class II fructose-bisphosphate aldolase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.

```

CC EMBL; U32734; AAC22182.1; -.
CC PIR; C64074; C64074.
CC HSSP; P11604; 1B57.
CC TIGR; H10524; -.
CC InterPro; IPR006411; Fruct bisp bact.
CC InterPro; IPR00771; K_bp_aldoase.
CC Pfam; PF0116; Fbp_aldoase; 1.
CC ProDom; PD002376; K_bp_aldoase; 1.
CC TIGRFAMs; TIGR0167; cBba; 1.
CC TIGRFAMs; TIGR01520; FruBisAldo II A; 1.
CC PROSITE; PS00602; ALDOLASE CLASS II_1; 1.
CC PROSITE; PS00806; ALDOLASE CLASS II_2; 1.
CC Lysase; Glycolysis; Zinc; Complete proteome.
KW METAL 108 108 ZINC (BY SIMILARITY).
FT METAL 111 111 ZINC (BY SIMILARITY).
SQ SEQUENCE 359 AA; 39339 MW; 1EDDFCD0B69E32C CRC64;

Query Match 39.1%; Score 43; DB 1; Length 359;
Best Local Similarity 70.0%; Pred. No. 26;
Matches 7; Conservative. 3; Mismatches 0; Indels 0; Gaps 0;

QY 7 ANYAGRGIRP 16
D 67 AFVAGKGIKP 76
|:|:|:|:|:|
|:|:|:|:|:|

RESULT 11
YAAJ_ECOLI
ID YAAJ_ECOLI STANDARD; PRT; 476 AA.
AC P30143;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative transporter yaaJ.
GN YAAJ OR B0007.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=92334977; PubMed=1630901;
RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
RA Isono K., Mizobuchi K., Nakata A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of
RT the 0-2.4 min region."
RL Nucleic Acids Res. 20:3305-3308 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474 (1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -!- SIMILARITY: BELONGS TO THE SODIUM:ALANINE SYMPORTER FAMILY
CC (SAF). STRONG, TO H. INFLUENZAE HI0189.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch.
```

```

CC EMBL; D10483; BAB96585.1; -.
CC EMBL; AE000111; AAC73118.1; -.
CC PIR; G64720; G64720.
CC EcoGene; EG11555; yaaJ.
CC InterPro; IPR002293; AA/rel_permease1.
CC InterPro; IPR001463; Na/Ala_symport.
CC Pfam; PF01235; Na_Ala_sym; 1.
CC PRINTS; PR00175; NAALASMPORT.
CC TIGRFAMs; TIGR00835; agcs; 1.
CC PROSITE; PS00873; NA ALANINE SYMP; 1.
KW Hypothetical protein; Transmembrane; Inner membrane; Transport;
KW Symport; Complete proteome.
FT TRANSMEM 4 24 POTENTIAL.
FT TRANSMEM 81 101 POTENTIAL.
FT TRANSMEM 141 161 POTENTIAL.
FT TRANSMEM 174 194 POTENTIAL.
FT TRANSMEM 207 227 POTENTIAL.
FT TRANSMEM 233 253 POTENTIAL.
FT TRANSMEM 300 320 POTENTIAL.
FT TRANSMEM 351 371 POTENTIAL.
FT TRANSMEM 391 411 POTENTIAL.
FT TRANSMEM 414 434 POTENTIAL.
SQ SEQUENCE 476 AA; 51662 MW; 2F6EB2E12E126E63 CRC64;

Query Match 39.1%; Score 43; DB 1; Length 476;
Best Local Similarity 44.4%; Pred. No. 34;
Matches 8; Conservative 2; Mismatches 2; Indels 6; Gaps 1;

QY 3 DIN-----PAWYAGRG1 14
D 120 DVNGQGRGPAWYMGRL 137
|:|:|:|:|:|
|:|:|:|:|:|

RESULT 12
CP12_CANFA
ID CP12_CANFA STANDARD; PRT; 511 AA.
AC P56592;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 1A2 (EC 1.14.14.1) (CYP1A2) (DAH2) (Cytochrome P450-
DE D2).
GN CYP1A2.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE OF 9-511 FROM N.A.
RC STRAIN=Beagle; TISSUE=Liver;
RX MEDLINE=91042464; PubMed=2122230;
RA Uchida T., Komori M., Kitada M., Kamataki T.;
RT "Isolation of cDNAs coding for three different forms of liver
RT microsomal cytochrome P-450 from polychlorinated biphenyl-treated
RT beagle dogs."
RL Mol. Pharmacol. 38:644-651 (1990).
RN [2]
RP SEQUENCE OF 1-16.
RC STRAIN=Beagle; TISSUE=Liver;
RX MEDLINE=89087526; PubMed=2910310;
RA Ohta K., Motoya M., Komori M., Miura T., Kitada M., Kamataki T.;
RT "A novel form of cytochrome P-450 in beagle dogs. P-450-D3 is a low
RT spin form of cytochrome P-450 but with catalytic and structural
RT properties similar to P-450d."
RL Biochem. Pharmacol. 38:91-96 (1989).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
```

CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -!- TISSUE SPECIFICITY: Constitutively expressed in liver.
 CC -!- INDUCTION: By polychlorinated biphenyl (PCB) in liver and kidney.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 DR PIR; B37222; B37222.
 DR HSPSP; P00179; LDT6.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR InterPro; IPR008066; EP4501a.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR01683; EP450ICYP1A.
 DR PRINTS; PR03385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum.
 FT INIT MET 0
 FT METAL 453 453 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 511 AA; 57505 MW; 200904C54F4B3CE7 CRC64;
 Query Match 39.1%; Score 43; DB 1; Length 511;
 Best Local Similarity 58.3%; Pred. No. 36;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 TPDINPAWVAGR 12
 DB 121 SPDSGPFVWAGR 132
 RESULT 13
 ID TRAZ2 DROVI STANDARD; PRT; 272 AA.
 AC C02008; C02009.
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Transformer-2 sex-determining protein.
 GN TRA-2.
 OS Drosophila virilis (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7244;
 [1]
 RN RP SEQUENCE FROM N.A. (ISOFORMS 272 AND 225).
 RX MEDLINE=97265426; PubMed=9111363;
 RA Chandler D., McGuffin M.E., Piskur J., Yao J., Baker B.S., Mattox W.;
 RT "Evolutionary conservation of regulatory strategies for the sex
 determination factor transformer-2.";
 RL Mol. Cell. Biol. 17:2808-2919(1997).
 CC -!- FUNCTION: Required for female sex determination in somatic cells
 and for spermatogenesis in male germ cells. Positive regulator of
 female-specific splicing and/or polyadenylation of doublesex (dsx)
 pre-mRNA. Splicing requires an enhancer complex. dsx repeat
 element, which contains six copies of a 13-nucleotide repeat and a
 purine-rich enhancer (PRE). DsxRE is formed through cooperative
 interactions between tra, tra2 and the sr proteins, and these
 interactions require both the repeat sequences and PRE. PRE is
 required for specific binding of tra2 to the dsxRE. Protein-RNA
 and protein-protein interactions are involved in tra-2 dependent
 activation and repression of alternative splicing (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=272;
 CC IsoId=002008-1; Sequence=Displayed;
 CC Name=225;
 CC IsoId=C02008-2; Sequence=VSP_005903;
 CC -!- DOMAIN: The RS2 (Arg/Ser-rich domain 2) and RNP-CS
 (ribonucleoprotein consensus sequence) domains are required for
 both male sterility and female-specific dsx splicing but the RS1
 domain is dispensable (By similarity).
 CC -!- PTM: Extensively phosphorylated on serine residues in the RS
 domain (By similarity).
 CC -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -!- SIMILARITY: Belongs to the SR family of splicing factors.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U72682; AAB58113.1; -;
 DR EMBL; U72682; AAB58112.1; -;
 DR FlyBase; FBgn0015686; Dvir\tra2.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; rrm; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS0102; RRM; 1.
 DR PROSITE; PS0030; RRM_RNP_1; 1.
 KW Sexual differentiation; Spermatogenesis; RNA-binding;
 KW Alternative splicing; Phosphorylation.
 FT DOMAIN 26 98 ARG/SER-RICH (RS1 DOMAIN).
 FT DOMAIN 105 183 RNA-BINDING (RRM).
 FT DOMAIN 184 204 LINKER.
 FT DOMAIN 205 272 ARG/SER-RICH (RS2 DOMAIN).
 FT VARSPLIC 3 49 Missing (in isoform 225).
 FT /FTID=VSP_005903.
 SQ SEQUENCE 272 AA; 31841 MW; 560DDA5DC3B26EE CRC64;
 Query Match 38.2%; Score 42; DB 1; Length 272;
 Best Local Similarity 52.6%; Pred. No. 29;
 Matches 10; Conservative 1; Mismatches 6; Indels 2; Gaps 1;
 QY 1 TPDINPAWVAGR 19
 DB 188 TP-TFGVNGRPSRPLGR 204
 RESULT 14
 ID ISPH MYCLE STANDARD; PRT; 332 AA.
 AC Q9X781.
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (EC 1.17.1.2).
 ISPH OR LYTB OR LYTB2 OR ML1938 OR MLCB1222.06C.
 OS Mycobacterium leprae.
 CC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 CC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 [1]
 RN RP SEQUENCE FROM N.A.
 RC STRAIN=TN;
 RX MEDLINE=21128732; PubMed=11234002;
 RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
 Wheeler P.A., Honore N., Garnier T., Churcher C., Harris D.,
 Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
 Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
 Holroyd S., Hornby T., Jagels K., Lacroix C., Maclean J., Moule S.,
 Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
 Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 Barrell B.G.;
 RA "Massive gene decay in the leprosy bacillus.";
 RL Nature 409:1007-1011(2001).
 CC -!- FUNCTION: Converts 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate
 into isopentenyl diphosphate (IPP) and dimethylallyl diphosphate
 (DMAPP) (By similarity).
 CC -!- CATALYTIC ACTIVITY: Isopentenyl diphosphate + NAD(P)(+) + H(2)O =
 CC (E)-4-hydroxy-3-methylbut-2-en-1-yl diphosphate + NAD(P)H.
 CC -!- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; seventh
 CC (last) step.
 CC -!- SIMILARITY: Belongs to the isph family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AL049491; CAB39812.1; ALT_INIT.
 DR EMBL; AL583923; CAB30893.1; ALT_INIT.
 DR HAMAP; MF01938; -.
 DR HAMAP; MF0191; -.
 DR InterPro; IPR003451; LytB.
 DR Pfam; PF02401; LYTB; 1.
 DR TIGRPFAMs; TIGR00216; isph_lytB; 1.
 KW Isoprene biosynthesis; Complete proteome; Oxidoreductase; NADP.
 SQ SEQUENCE 332 AA; 36225 MW; F2A92A06EC35B13 CRC64;

Query Match 38.2%; Score 42; DB 1; Length 332;
 Best Local Similarity 77.8%; Pred. No. 34;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 DINDAWYAG 11
 Db 270 DIDPAWLAG 278

RESULT 15

YHHJ_ECOLI STANDARD; PRT; 374 AA.
 ID YHHJ_ECOLI
 AC P31993; P76703;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Hypothetical protein yhhj.
 GN YHHJ OR B3485 OR SF3501 OR S4262.
 OS Escherichia coli, and
 OS Shigella flexneri.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562, 623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=KI2 / MG1655;
 RX MEDLINE=94316500; PubMed=8041620;
 RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
 RT "Analysis of the Escherichia coli genome. V. DNA sequence of the
 RT region from 76.0 to 81.5 minutes."
 RL Nucleic Acids Res. 22:2576-2586(1994).
 RN [2]
 RP SEQUENCE OF 225-374 FROM N.A.
 RC SPECIES=E.coli; STRAIN=KI2;
 RX MEDLINE=93259920; PubMed=8387990;
 RA Zhao S., Sandt C.H., Feulner G., Vlazny D.A., Gray J.A., Hill C.W.;
 RT "Rhs elements of Escherichia coli K-12: complex composites of shared
 RT and unique components that have different evolutionary histories."
 RL J. Bacteriol. 175:2799-2808(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
 RX MEDLINE=22272406; PubMed=12384590;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
 RT through comparison with genomes of Escherichia coli K12 and O157";
 RL Nucleic Acids Res. 30:4432-4441(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
 RX MEDLINE=22590274; PubMed=12704152;

RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
 RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
 RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
 RA Schwartz D.C., Blattner F.R.;
 RT "Complete genome sequence and comparative genomics of Shigella
 RT flexneri serotype 2a strain 2457T";
 RL Infect. Immun. 71:2775-2786(2003).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (Potential).
 CC -!- SIMILARITY: Belongs to the ABC-2 integral membrane protein family.
 CC -!- SIMILARITY: TO E.COLI YBHR AND YBHS AND TO THE C-TERMINAL OF
 CC E.COLI YHIG.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; U00039; AAB18460.1; ALT_INIT.
 DR EMBL; AE000424; AAC76510.1; ALT_INIT.
 DR EMBL; L02370; AAC61886.1; -.
 DR EMBL; AE015358; AAN44960.1; ALT_INIT.
 DR EMBL; AE016932; AAP19222.1; -.
 DR EcoGene; EGI1767; yhhj.
 DR InterPro; IPR000412; ABC_transpt2.
 DR PROSITE; PS00890; ABC2_MEMBRANE; 1.
 KW Hypothetical protein; Transport; Transmembrane; Inner membrane;
 KW Complete proteome.
 FT TRANSMEM 23 43 POTENTIAL.
 FT TRANSMEM 173 193 POTENTIAL.
 FT TRANSMEM 230 250 POTENTIAL.
 FT TRANSMEM 256 276 POTENTIAL.
 FT TRANSMEM 284 304 POTENTIAL.
 FT TRANSMEM 343 363 POTENTIAL.
 SQ SEQUENCE 374 AA; 41061 MW; 02895FBI3F493391 CRC64;

Query Match 38.2%; Score 42; DB 1; Length 374;
 Best Local Similarity 50.0%; Pred. No. 38;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PDINPAWYAG 11
 Db 169 PNLDPWAFGG 178

Search completed: August 12, 2004, 14:44:54
 Job time : 6.23256 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 27.5581 seconds
(without alignments)
228.984 Million cell updates/sec

Title: US-09-700-643A-12

Perfect score: 110

Sequence: 1 TPDINPANYAGRGIRPVGR 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_prodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109	99.1	98	Q8WN12	Q8WN12 ovis aries
2	84	76.4	117	Q9W624	Q9W624 carassius a
3	57	51.8	54	Q7Z6Y1	Q7Z6Y1 homo sapien
4	57	51.8	465	O60687	O60687 homo sapien
5	57	51.8	465	Q8WN85	Q8WN85 homo sapien
6	54	49.1	1120	Q7UPG7	Q7UPG7 rhodopirell
7	53	48.2	467	Q8R054	Q8R054 mus musculus
8	53	48.2	468	Q8X1F8	Q8X1F8 mus musculus
9	52	47.3	419	Q8UKP2	Q8UKP2 agrobacteri
10	51	46.4	333	Q96SD4	Q96SD4 homo sapien
11	51	46.4	367	Q8NSQ2	Q8NSQ2 homo sapien
12	51	46.4	367	Q8IXT2	Q8IXT2 homo sapien
13	51	46.4	370	Q8CGW9	Q8CGW9 mus musculus
14	51	46.4	430	Q27142	Q27142 methanobact
15	50	45.5	130	Q82RV0	Q82RV0 streptomyce
16	50	45.5	527	Q76383	Q76383 caenorhabdi

17	49	44.5	790	10	Q9M371	Q9M371 arabidopsis
18	49	44.5	1654	16	Q8FM04	Q8FM04 corynebacte
19	48.5	44.1	664	16	Q911W4	Q911W4 pseudomonas
20	48	43.6	343	3	O74569	O74569 coprinus ci
21	48	43.6	402	16	Q821E7	Q821E7 streptomyce
22	47	42.7	540	16	Q886T3	Q886T3 pseudomonas
23	46.5	42.3	458	16	Q8FRJ9	Q8FRJ9 corynebacte
24	46	41.8	144	4	Q96MP2	Q96MP2 homo sapien
25	46	41.8	299	16	Q87F82	Q87F82 xylella fas
26	46	41.8	333	16	Q8PH76	Q8PH76 xylella fas
27	46	41.8	369	16	Q98MB6	Q98MB6 rhizobium l
28	46	41.8	369	16	Q8G4X2	Q8G4X2 bifidobacte
29	46	41.8	424	2	Q9EUJ0	Q9EUJ0 salmonella
30	46	41.8	424	16	Q93GQ6	Q93GQ6 salmonella
31	46	41.8	439	16	Q89IP7	Q89IP7 bradyrhizob
32	46	41.8	820	5	Q9NE93	Q9NE93 leishmania
33	45.5	41.4	333	16	Q9RJ10	Q9RJ10 streptomyce
34	45.5	41.4	425	16	Q986U6	Q986U6 rhizobium l
35	45.5	41.4	779	3	Q9P5J9	Q9P5J9 neurospora
36	45	40.9	179	5	Q20170	Q20170 caenorhabdi
37	45	40.9	187	16	Q7URC6	Q7URC6 rhodopirell
38	45	40.9	216	17	Q8ZU76	Q8ZU76 pyrobaculum
39	45	40.9	226	16	Q9A359	Q9A359 caulobacter
40	45	40.9	230	16	Q88H95	Q88H95 pseudomonas
41	45	40.9	284	17	O50128	O50128 pyrococcus
42	45	40.9	284	17	Q8U112	Q8U112 pyrococcus
43	45	40.9	419	13	Q7ZTL7	Q7ZTL7 xenopus lae
44	45	40.9	424	16	Q8ZIG9	Q8ZIG9 salmonella
45	45	40.9	472	12	Q91FX2	Q91FX2 chilo iride

ALIGNMENTS

RESULT 1

Q8WN12 PRELIMINARY; PRT; 98 AA.

ID Q8WN12
AC Q8WN12;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DE 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Preprolactin-releasing peptide.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID:9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Curlewis J.D.; Kusters D.H.L.; Barclay J.L.; Anderson S.T.;
RT "Prolactin-releasing peptide (prp) in the ewe: cDNA cloning, mRNA
RT distribution and effects on prolactin secretion in vitro and in
RT vivo."
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF450453; AAL47178.1;
SQ SEQUENCE 98 AA; 10513 MW; 2A53331ED62CAAB5 CRC64;

Query Match 99.1%; Score 109; DB 6; Length 98;
Best Local Similarity 100.0%; Pred.No. 9.7e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPANYAGRGIRPVGR 19

Db 34 TPDINPANYAGRGIRPVGR 52

RESULT 2

Q9W624 PRELIMINARY; PRT; 117 AA.

ID Q9W624
AC Q9W624;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)


```

OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294145; CAD75095.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1120 AA; 123221 MW; 67C757B73421BF9B CRC64;

Query Match 49.1%; Score 54; DB 16; Length 1120;
Best Local Similarity 64.3%; Pred. No. 21;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 DINPAWYAGRGIRP 16
Db :|:|:|:|:|:|
202 EISPAWAKWGIRP 215

RESULT 7
QY Q8R054 PRELIMINARY; PRT; 467 AA.
ID Q8R054;
AC Q8R054;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Similar to sushi-repeat protein.
GN 1110039C07RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC028307; AAH28307.1; -.
DR MGD; MGI:1916042; 1110039C07RIK.
CO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR003410; Hvalin.
DR Pfam; PF02494; HYR; 1.
DR SMART; SM00084; sush1; 3.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
SQ SEQUENCE 467 AA; 53009 MW; BB4C01C7E6118BE0 CRC64;

Query Match 48.2%; Score 53; DB 11; Length 467;
Best Local Similarity 53.3%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 PDINPAWYAGRGIRP 16
Db :|:|:|:|:|:|
19 PAVTPTWYAGSGVSP 33

RESULT 8
QY Q8K1F8 PRELIMINARY; PRT; 468 AA.
ID Q8K1F8;
AC Q8K1F8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Sushi-repeat containing protein.
GN 1110039C07RIK.

```

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Huang C.-H., Chen H., Peng J., Chen Y.;
RT "Cloning and characterization of the sushi-repeat containing protein
RT (SRP) as a novel interaction partner of Rh type C glycoprotein
RT (RhCG).";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF393640; AAM73691.1; -.
DR MGD; MGI:1916042; 1110039C07RIK.
CO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR003410; Hvalin.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF02494; HYR; 1.
DR Pfam; PF00084; sush1; 3.
DR SMART; SM00032; CCP; 3.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
SQ SEQUENCE 468 AA; 53180 MW; 151A952070D040D4 CRC64;

Query Match 48.2%; Score 53; DB 11; Length 468;
Best Local Similarity 53.3%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 PDINPAWYAGRGIRP 16
Db :|:|:|:|:|:|
20 PAVTPTWYAGSGVSP 34

RESULT 9
QY Q8UKP2 PRELIMINARY; PRT; 419 AA.
ID Q8UKP2;
AC Q8UKP2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ABC transporter, substrate binding protein.
GN UGPB OR AU05058 OR AGR_FAT_77.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OG Plasmid AT.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chunley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RA "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Hailing C., Mullin L.,
RA Houmel K., Gordon J., Vaudin M., Iatchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursun J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RA "Genome sequence of the plant pathogen and biotechnology agent
RT

```

```

Query Match          46.4%; Score 51; DB 4; Length 367;
Best Local Similarity 37.5%; Pred. No. 19;
Matches 12; Conservative 3; Mismatches 3; Indels 14; Gaps 0
QY      1 TPD I - N P A W Y A ----- G G I R P V G 18
      ||| : |||| : |||| : |||| : |||| : |||| :
DB      311 T P S V F P N P A W I S L L H F C G P P A G G R F Q P V G 342

RESULT 13
Q8CGW3 Q8CGW3 PRELIMINARY; PRT; 370 AA.
AC Q8CGW9
ID Q8CGW9;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DD 01-MAR-2003 (TREMBLrel. 23, Last sequence update)

```



```

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Doublesex and mab-3 related transcription factor 7.
GN DMRTC2 OR 493343221R1K OR DMYR17.
OS Mus musculus (Mouse).
OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA Kim S., Kettlewell J., Anderson R.C., Bardwell V.J., Zarkower D.;
RT "Sexually dimorphic expression of multiple doublesex-related genes in
RT the embryonic gonad.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF542046; AN77233.1; -.
DR MGD; MGI:1918491; Dmrtc2.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0007548; P:sex differentiation; IEA.
DR InterPro; IPR001275; DM-DNA-binding.
DR Pfam; PF00751; DM-domain; 1.
DR SMART; SM00301; DM; 1.
DR PROSITE; PS40000; DM_DOMAIN_1; 1.
DR PROSITE; PS50809; DM_DOMAIN_2; 1.
SQ SEQUENCE 370 AA; 33095 MW; F4DFCB23C94C8A6C CRC64;

Query Match 46.4%; Score 51; DB 11; Length 370;
Best Local Similarity 37.5%; Pred. No. 19;
Matches 12; Conservative 3; Mismatches 3; Indels 14; Gaps 2;

QY 1 TPDI-1NPAWYA-----GGRGIRPVG 18
DB 314 TPSVPPNPAMISLLHCPGPPGGRGIRPVG 345

RESULT 14
ID 027142 PRELIMINARY; PRT; 430 AA.
AC
CQ 027142;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Conserved protein.
GN MTH1070.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiawani N., Caruso A., Bush D., Safer H., Fatwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
DR EMBL; AE000877; AAE85559.1; -.

DR PIR; B69009; B69009.
DR InterPro; IPR002510; Peptidase_U52.
DR Pfam; PF01523; FmbA_TiD; 1.
KW Complete proteome.
SQ SEQUENCE 430 AA; 46062 MW; C4FAE47C111749B6 CRC64;

Query Match 46.4%; Score 51; DB 17; Length 430;
Best Local Similarity 69.2%; Pred. No. 22;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 DINPAWVAGRGIR 15
DB 191 DINFEWVAGRACR 203

RESULT 15
Q82RV0
ID Q82RV0 PRELIMINARY; PRT; 130 AA.
AC
CQ 082RV0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative transposase.
GN SAV43.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:528-531(2003).
DR EMBL; AP005021; BAC67752.1; -.
KW Complete proteome.
SQ SEQUENCE 130 AA; 14603 MW; 1642339D0C67827E CRC64;

Query Match 45.5%; Score 50; DB 16; Length 130;
Best Local Similarity 69.2%; Pred. No. 9;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 PAWVAGRGIRPVG 18
DB 33 PAWLAGRGQPEG 45

Search completed: August 12, 2004, 14:49:06
Job time : 27.5581 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 42.4419 Seconds
(without alignments)
133.146 Million cell updates/sec

Title: US-09-700-643A-12

Perfect score: 110
Sequence: 1 TPDINPAVAGRIIPVGRX 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	109	99.1	19	2 AAW31370	Aaw31370 Bovine G
2	109	99.1	19	2 AAW95185	Aaw95185 Bovine pi
3	109	99.1	19	6 ABU60830	Abu60830 Peptide p
4	109	99.1	20	2 AAW31374	Aaw31374 Bovine G
5	109	99.1	20	2 AAW95191	Aaw95191 Bovine pi
6	109	99.1	20	2 AAW97232	Aaw97232 Bovine pi
7	109	99.1	20	3 AAY49301	Aay49301 19p2 liga
8	109	99.1	20	3 AAB10350	Aab10350 Bovine ox
9	109	99.1	20	4 AAB90996	Aab90996 Prolactin
10	109	99.1	20	4 AAB46954	Aab46954 Peptide p
11	109	99.1	20	4 AAG62519	Aag62519 Bovine CR
12	109	99.1	20	5 AAE26402	Aae26402 Bovine Pr
13	109	99.1	20	6 ABU60834	Abu60834 Peptide p
14	109	99.1	21	2 AAW31375	Aaw31375 Bovine G
15	109	99.1	21	2 AAW95192	Aaw95192 Bovine pi
16	109	99.1	21	3 AAB10351	Aab10351 Bovine ox
17	109	99.1	21	4 AAG62520	Aag62520 Bovine CR
18	109	99.1	21	6 ABU60835	Abu60835 Peptide p
19	109	99.1	22	2 AAW31376	Aaw31376 Bovine G
20	109	99.1	22	2 AAW95193	Aaw95193 Bovine pi
21	109	99.1	22	3 AAB10352	Aab10352 Bovine ox
22	109	99.1	22	4 AAG62521	Aag62521 Bovine CR
23	109	99.1	22	6 ABU60836	Abu60836 Peptide p
24	109	99.1	31	2 AAW31371	Aaw31371 Bovine G
25	109	99.1	31	2 AAW95188	Aaw95188 Bovine pi

26	109	99.1	31	2 AAW97613	Aaw97613 Bovine 19
27	109	99.1	31	2 AAW97218	Aaw97218 Bovine pi
28	109	99.1	31	3 AAY49298	Aay49298 19p2 liga
29	109	99.1	31	3 AAY49290	Aay49290 19p2 liga
30	109	99.1	31	3 AAB10347	Aab10347 Bovine ox
31	109	99.1	31	4 AAG62516	Aag62516 Bovine CR
32	109	99.1	31	5 AAE26399	Aae26399 Bovine Pr
33	109	99.1	31	6 ABU60825	Abu60825 Peptide p
34	109	99.1	31	6 ABU60831	Abu60831 Peptide p
35	109	99.1	32	2 AAW31372	Aaw31372 Bovine G
36	109	99.1	32	2 AAW95189	Aaw95189 Bovine pi
37	109	99.1	32	3 AAB10348	Aab10348 Bovine ox
38	109	99.1	32	4 AAG62517	Aag62517 Bovine CR
39	109	99.1	32	6 ABU60832	Abu60832 Peptide p
40	109	99.1	33	2 AAW31373	Aaw31373 Bovine G
41	109	99.1	33	2 AAW95190	Aaw95190 Bovine pi
42	109	99.1	33	3 AAY49297	Aay49297 19p2 liga
43	109	99.1	33	3 AAB10349	Aab10349 Bovine ox
44	109	99.1	33	4 AAG62518	Aag62518 Bovine CR
45	109	99.1	33	6 ABU60833	Abu60833 Peptide p

ALIGNMENTS

RESULT 1
AAW31370
ID AAW31370 standard; peptide, 19 AA.
XX
AC AAW31370;
XX
DT 06-APR-1998 (first entry)
XX
DE Bovine G protein-coupled receptor ligand fragment from P-2 fraction.
XX
KW G protein-coupled receptor; ligand binding; pharmaceutical; modulator;
KW Pituitary; central nervous system; pancreas; prophylactic;
KW Therapeutic agent.
XX
OS Bos taurus.
XX
PN WO9724436-A2.
XX
PD 10-JUL-1997.
XX
PF 26-DEC-1996; 96WO-JP003821.
XX
PR 28-DEC-1995; 95JP-00343371.
PR 15-MAR-1996; 96JP-00059419.
PR 12-AUG-1986; 86JP-00211805.
PR 18-SEP-1996; 96JP-00246573.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Hinuma S, Habata Y, Kawanata Y, Hosoya M, Fujii R, Fukusumi S;
PI Kitada C;
XX
PI WPI; 1997-363672/33.
DR N-PSDB; AAV02393.
XX
PT Ligand peptide for G protein-coupled receptor - acts by modulating
PT function in the central nervous system, pancreas and pituitary gland.
XX
PS Claim 2; Page 160; 258pp; English.
CC This sequence represents a peptide fragment of a novel bovine pituitary
CC derived ligand corresponding to amino acid residues 34 to 52 of the
CC sequence in AAW31368. This fragment was obtained by purification and
CC analysis of the N-terminal sequence of a P-2 fraction which is used in an
CC assay to monitor ligand binding to the G protein-coupled receptor
CC protein. Pharmaceutical compositions containing this ligand may be used
CC as a pituitary function modulator, a central nervous system modulator or
CC a pancreatic function modulator. This ligand could have specific

CC applications as a prophylactic or therapeutic agent for dementia,
 CC depression, hyperkinetic syndrome, disturbance of consciousness, anxiety
 CC syndrome, schizophrenia, trauma, growth hormone secretory disease, hyper-
 CC and polyphagia, hypercholesterolaemia, hyperglycaemia,
 CC hyperlipidaemia, hyperprolactinaemia, diabetes, cancer, pancreatitis,
 CC renal disease, Turner's syndrome, neurosis, rheumatoid arthritis, spinal
 CC injury, transient brain ischaemia, amyotrophic lateral sclerosis, acute
 CC myocardial infarction, spinocerebellar degeneration, bone fracture,
 CC trauma, atopic dermatitis, osteoporosis, asthma, epilepsy, infertility
 CC and/or oligogalactia. Assays can also be developed to screen compounds
 CC which are capable of altering the binding activity of the ligand thus
 CC affecting activation of the G protein-coupled receptor protein
 XX
 SQ Sequence 19 AA;

Query Match 99.1%; Score 109; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 2e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
 |||||
 DB 1 TPDINPAWYAGRGIRPVGR 19

RESULT 2
 AAW95185
 ID AAW95185 standard; peptide; 19 AA.
 AC AAW95185;
 XX
 XX
 DT 10-MAR-1999 (first entry)
 XX

DE Bovine pituitary-derived ligand polypeptide partial sequence.

XX Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
 KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
 KW Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
 KW gene therapy; transgenic animal; bovine.
 XX

OS Bos sp.
 XX
 XX WO9849295-A1.
 PN
 XX
 PD 05-NOV-1998.
 XX
 XX 27-APR-1998; 98WO-JP001923.
 XX
 XX 28-APR-1997; 97JP-00109974.
 XX
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX

PI Hinuma S, Fukusumi S;
 XX
 XX WPI; 1999-009423/01.
 DR
 XX

PT New polypeptide ligand for orphan G protein coupled receptors - used for
 PT treating disorders of central nervous system, pituitary and pancreas, and
 PT for drug screening.
 XX

PS Example 16; Page 145; 205pp; English.

XX The invention relates to a murine pituitary-derived ligand polypeptide
 CC which is a ligand for the G-protein coupled orphan receptor designated
 CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing
 CC the ligand polypeptide encoding DNA are used to produce a recombinant
 CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate
 CC function of the pituitary, central nervous system, pancreas and other
 CC tissues and can be used to screen for agents that modulate binding of the
 CC polypeptide to the receptor; to quantify the amount of receptor in a
 CC sample and to raise antibodies. They may also be used therapeutically.

CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's
 CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;
 CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;
 CC rheumatoid arthritis; epilepsy and many others, also to improve post-
 CC operative nutritional status and as vasopressor. Transgenic animals
 CC carrying the ligand polypeptide encoding DNA or its mutin are used to
 CC study the function of the polypeptide-expressing genes, as models of
 CC disease, for drug screening and as source of cell lines. The ligand
 CC polypeptide DNA is used as a source of probes and primers; to identify
 CC related sequences; in receptor-binding assays; for production of Ab and
 CC antisera; in drug development; for gene therapy and to develop transgenic
 CC animals. The present sequence represents a bovine pituitary-derived
 CC ligand polypeptide obtained by N-terminal sequence analysis
 XX
 SQ Sequence 19 AA;

Query Match 99.1%; Score 109; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 2e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
 |||||
 DB 1 TPDINPAWYAGRGIRPVGR 19

RESULT 3
 ABU60830
 ID ABU60830 standard; peptide; 19 AA.
 AC ABU60830;
 XX

DT 06-MAY-2003 (first entry)
 XX

DE Peptide production by gene recombination associated peptide #14.

KW Peptide production; low-molecular peptide; KISS-1; GPR8 ligand;
 KW gene recombination.
 XX

OS Bos sp.

XX WO200292829-A1.
 PN

XX 21-NOV-2002.
 PD

XX 16-MAY-2002; 2002WO-JP004735.
 PF

XX 17-MAY-2001; 2001JP-00147341.
 PR

XX (TAKE) TAKEDA CHEM IND LTD.
 PA

XX Nishimura O, Suenaga M, Ito T, Kitada C;
 PI

XX WPI; 2003-129302/12.
 DR

XX Process for producing peptides e.g. Kiss-1 peptide and GPR8 ligand for
 XX subsequent applications by gene recombination technique through tandem
 XX repeats to provide precursor protein with specific cleavage sites.
 PT

PS Disclosure; Page 60; 87pp; Japanese.

XX The invention describes a method of producing a peptide comprising the
 CC excision of the N and C-terminals of a target peptide with enzymes or
 CC chemically through the attached cleavage sites repeated by ligation in a
 CC precursor protein. The method is for producing (low-molecular) peptides
 CC e.g. Kiss-1 peptide and GPR8 ligand for subsequent applications by the
 CC gene recombination technique through tandem repeats to provide a
 CC precursor protein with specific cleavage sites. With this method, peptide
 CC production can be carried out easily to provide large quantities of the
 CC required peptides. This is the amino acid sequence of a peptide
 CC associated with the peptide production method of the invention
 XX

SQ Sequence 19 AA;

Query Match 99.1%; Score 109; DB 6; Length 19;
 Best Local Similarity 100.0%; Pred. No. 2e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
 |||||
 DB 1 TPDINPAWYAGRGIRPVGR 19
 |||||

RESULT 4
 AAW31374
 ID AAW31374 standard; peptide; 20 AA.
 AC AAW31374;
 DT 06-APR-1998 (first entry)
 DE Bovine G protein-coupled receptor ligand peptide fragment 4.
 KW G protein-coupled receptor; ligand binding; pharmaceutical; modulator;
 KW pituitary; central nervous system; pancreas; prophylactic;
 KW therapeutic agent.
 XX
 OS Bos taurus.
 XX
 PN WO9724436-A2.
 XX
 PD 10-JUL-1997.
 XX
 PF 26-DEC-1996; 96WO-JF003821.
 XX
 PR 28-DEC-1995; 95JP-00343371.
 PR 15-MAR-1996; 96JP-00059419.
 PR 12-AUG-1996; 96JP-00211805.
 PR 18-SEP-1996; 96JP-00246573.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;
 PI Kitada C;
 XX
 WPI; 1997-363672/33.
 DR N-PSDB; AAV02397.
 XX
 PT Ligand peptide for G protein-coupled receptor - acts by modulating
 PT function in the central nervous system, pancreas and pituitary gland.
 PS
 PS Claim 2; Page 161; 258pp; English.
 XX
 CC This sequence represents a peptide fragment of a novel bovine pituitary
 CC derived ligand corresponding to amino acid residues 34 to 53 of the
 CC sequence in AAW31368 and is used in an assay to monitor ligand binding to
 CC the G protein-coupled receptor protein. Pharmaceutical compositions
 CC containing this ligand may be used as a pituitary function modulator, a
 CC central nervous system modulator or a pancreatic function modulator. This
 CC agent could have specific applications as a prophylactic or therapeutic
 CC ligand for dementia, depression, hyperkinetic syndrome, disturbance of
 CC consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone
 CC secretory disease, hyper and polyphagia, hypercholesterolaemia,
 CC hyperglycaemia, hyperlipidaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia,
 CC ankylosing lateral sclerosis, acute myocardial infarction,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis, asthma, epilepsy, infertility and/or oligogalactia. Assays
 CC can also be developed to screen compounds which are capable of altering
 CC the binding activity of the ligand thus affecting activation of the G
 CC protein-coupled receptor protein
 XX
 SQ Sequence 20 AA;

Query Match 99.1%; Score 109; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
 |||||
 DB 1 TPDINPAWYAGRGIRPVGR 19
 |||||

RESULT 5
 AAW95191
 ID AAW95191 standard; peptide; 20 AA.
 AC AAW95191;
 XX
 DT 10-MAR-1999 (first entry)
 DE Bovine pituitary-derived ligand polypeptide fragment.
 DE XX
 KW Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
 KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
 KW Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
 KW gene therapy; transgenic animal; bovine.
 XX
 OS Bos sp.
 XX
 PN WO9849295-A1.
 XX
 PD 05-NOV-1998.
 XX
 PF 27-APR-1998; 98WO-JP001923.
 XX
 PR 28-APR-1997; 97JP-00109974.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Hinuma S, Fukusumi S;
 XX
 WPI; 1999-009423/01.
 DR
 PT New polypeptide ligand for orphan G protein coupled receptors - used for
 PT treating disorders of central nervous system, pituitary and pancreas, and
 PT for drug screening.
 PS
 PS Example 19; Page 151; 206pp; English.
 XX
 CC The invention relates to a murine pituitary-derived ligand polypeptide
 CC which is a ligand for the G-protein coupled orphan receptor designated
 CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing
 CC the ligand polypeptide encoding DNA are used to produce a recombinant
 CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate
 CC function of the pituitary, central nervous system, pancreas and other
 CC tissues and can be used to screen for agents that modulate binding of the
 CC polypeptide to the receptor; to quantify the amount of receptor in a
 CC sample and to raise antibodies. They may also be used therapeutically,
 CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's
 CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;
 CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;
 CC rheumatoid arthritis, epilepsy and many others, also to improve post-
 CC operative nutritional status and as vasopressor. Transgenic animals
 CC carrying the ligand polypeptide encoding DNA or its mutin are used to
 CC study the function of the polypeptide-expressing genes, as models of
 CC disease, for drug screening and as source of cell lines. The ligand
 CC polypeptide DNA is used as a source of probes and primers; to identify
 CC related sequences; in receptor-binding assays; for production of Ab and
 CC antisera; in drug development; for gene therapy and to develop transgenic
 CC animals. The present sequence represents a bovine genome-derived ligand
 CC polypeptide fragment which is similar to the murine ligand polypeptide
 XX
 SQ Sequence 20 AA;

Query Match 99.1%; Score 109; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
| | | | | | | | | | | | | | | | | | | | |
Db 1 TPDINPAWYAGRGIRPVGR 19

RESULT 6

AAW97232
ID AAW97232 standard; peptide; 20 AA.

XX AC AAW97232;

XX DT 06-MAY-1999 (first entry)

XX DE Bovine pituitary-derived ligand polypeptide fragment.

XX KW Bovine pituitary-derived ligand; modulation; prolactin secretion;
XX KW G protein-coupled receptor; GPCR; hypocoarixianism; gonocyst cacosgenesis;
XX KW menopausal syndrome; euthyroid; hypometabolism; lactation;
XX KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;
XX KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;
XX KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;
XX KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;
XX KW contraceptive; placental function; choriocarcinoma; hydatid mole;
XX KW interruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;
XX KW abnormal lipidmetabolism; oxytocia.

XX OS Bos sp.

XX PN WO9858962-A1.

XX PD 30-DEC-1998.

XX PF 22-JUN-1998; 98WO-JP002765.

XX PR 23-JUN-1997; 97JP-00165437.

XX PA (TAKE) TAKEDA CHEM IND LTD.

XX PI Hinuma S, Kawamata Y, Fujii R, Matsumoto H;

XX PS WPI; 1999-105614/09.

XX PT Use of G protein-coupled receptor ligands - for modulating prolactin
XX PT secretion or placental function, e.g. for treating menopausal syndrome,
XX PT tumours, autoimmune disease or abnormal pregnancy.

XX PS Claim 3; Page 136; 241pp; English.

XX CC The present sequence represents a bovine pituitary-derived ligand
XX CC fragment. It is used in the course of the invention. The specification
XX CC describes an agent for modulating prolactin secretion which comprises a
XX CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)
XX CC protein. The agents for promoting prolactin secretion can be used for
XX CC treating or preventing hypocoarixianism, gonocyst cacosgenesis, menopausal
XX CC syndrome, euthyroid or hypometabolism. They can be used for promoting
XX CC lactation in a domestic mammal and as an aphrodisiac. The agents for
XX CC inhibiting prolactin secretion can be used for treating or preventing
XX CC pituitary adenomatosis, brain tumour, emmenopathy, autoimmune disease,
XX CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,
XX CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome, Forbes-
XX CC Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia. The
XX CC inhibitory agents can also be used as contraceptives. The agents for
XX CC modulating placental function can be used for treating or preventing
XX CC choriocarcinoma, hydatid mole, interruption mole, abortion, unthrifty fetus,
XX CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia

XX SQ Sequence 20 AA;

Query Match 99.1%; Score 109; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
| | | | | | | | | | | | | | | | | | | | |
Db 1 TPDINPAWYAGRGIRPVGR 19

RESULT 7

AA49301
ID AA49301 standard; peptide; 20 AA.

XX AC AA49301;

XX DT 22-FEB-2000 (first entry)

XX DE 19P2 ligand peptide fragment.

XX KW Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;
XX KW pituitary; regulatory mechanism; central nervous system; pancreatic.

XX OS Bos sp.

XX FT Key Location/Qualifiers
XX FT Modified-site 20
XX FT /note= "C-terminal amide"

XX PN WO9960112-A1.

XX PD 25-NOV-1999.

XX PF 20-MAY-1999; 99WO-JP002650.

XX PR 21-MAY-1998; 98JP-00140293.

XX PA (TAKE) TAKEDA CHEM IND LTD.

XX PI Matsumoto H, Kitada C, Hinuma S;

XX PS WPI; 2000-039381/03.

XX PT New monoclonal antibodies, useful in diagnosis, as drugs and in studying
XX PT diseases related to ligand abnormality.

XX PS Disclosure; Page 27; 73pp; Japanese.

XX CC The invention provides a monoclonal antibody which has a specific
XX CC reaction with the part peptide of the C-terminal of 19P2 ligand or its
XX CC derivative. The antibodies can be used in diagnosis or to treat or
XX CC prevent diseases associated with abnormality in the pituitary function
XX CC regulatory mechanism (e.g. promotion of prolactin secretion), central
XX CC nervous regulatory mechanism, and pancreatic function regulatory
XX CC mechanism. The antibody-based immunoassay can also be applied in
XX CC clarifying the physiological functions of the ligand and its derivative.

XX SQ Sequence 20 AA;

Query Match 99.1%; Score 109; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
| | | | | | | | | | | | | | | | | | | | |
Db 1 TPDINPAWYAGRGIRPVGR 19

RESULT 8

AAB10350
ID AAB10350 standard; peptide; 20 AA.

XX AC AAB10350;

XX DT 24-NOV-2000 (first entry)

XX Bovine oxytocin secretion promoting peptide SEQ ID NO: 6.
DE
XX Bovine; oxytocin secretion promoter; G protein-coupled receptor protein;
XX treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
KW caesarean section; artificial fertilization; galactostasis; goat; pig;
KW veterinary medicine; milk production.
XX
XX Bos taurus.
OS
XX WO200039704-A1.
XX
XX 06-JUL-2000.
XX
XX 22-DEC-1999; 99WO-JP007199.
XX
XX 25-DEC-1998; 98JP-00369585.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Matsumoto H., Kitada C, Hinuma S;
XX
XX WPI; 2000-452298/39.
XX
XX Physiologically-active polypeptide recognized as ligand by G protein-
XX coupled receptor protein, for promoting secretion of oxytocin, as drugs
XX for diseases relating to oxytocin secretion and in veterinary medicine.
XX
XX Claim 5; Page 51; 72pp; Japanese.
XX
XX This invention describes a novel oxytocin secretion-regulating agent
XX which contains a ligand peptide or its salt for the G protein-coupled
XX receptor protein. It is useful in the form of drugs for ameliorating,
XX preventing and treating diseases relating to oxytocin secretion e.g. weak
XX pains and atonic bleeding, before and after expulsion of placenta,
XX uterine recovery failure, caesarean section, stoppage of artificial
XX fertilization or galactostasis and is also applicable in veterinary
XX medicine for promoting milk production in cow, goat and pig. This
XX sequence represents a bovine peptide which acts as an oxytocin secretion
XX promoter
XX
XX Sequence 20 AA;
XX
XX Query Match 99.1%; Score 109; DB 3; Length 20;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-10;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 TPDINPAWYAGRGIRPVGR 19
XX | | | | | | | | | | | | | | | | | | | | | |
XX Db 1 TPDINPAWYAGRGIRPVGR 19
XX
XX RESULT 9
XX AAB90996
XX ID AAB90996 standard; peptide; 20 AA.
XX
XX AC AAB90996;
XX
XX 22-JUN-2001 (first entry)
XX
XX Prolactin releasing peptide SEQ ID NO:170.
XX
XX Protection; endogenous therapeutic peptide; peptidase; conjugation;
XX blood component; modification; succinimide; maleimido group; amino;
XX hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
XX Homo sapiens.
XX Synthetic.
XX
XX WO200069900-A2.
XX
XX 23-NOV-2000.
XX
XX

PF 17-MAY-2000; 2000WO-US013576.
XX
XX 17-MAY-1999; 99US-0134406P.
XX
XX 10-SEP-1999; 99US-0153406P.
XX
XX 15-OCT-1999; 99US-0159783P.
XX
XX (CONJ-) CONJUCHEM INC.
XX
XX Bridon DF, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;
XX
XX WPI; 2001-112059/12.
XX
XX Modifying and attaching therapeutic peptides to albumin prevents
XX peptidase degradation, useful for increasing length of in vivo activity.
XX
XX Disclosure; Page 245; 733pp; English.
XX
XX The present invention describes a modified therapeutic peptide (I)
XX comprising a therapeutically active amino acid region (III) and a
XX reactive group (II) (e.g. succinimide and maleimido groups) attached to
XX a less therapeutically active amino acid region (IV), which covalently
XX bonds with amino/hydroxyl/thiol groups on blood components to form a
XX peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
XX (I) are useful for modifying therapeutic peptides e.g. hormones, growth
XX factors and neurotransmitters, to protect them from peptidase activity in
XX vivo for the treatment of various disorders. Endogenous therapeutic
XX peptides are not suitable as drug candidates as they require frequent
XX administration due to rapid degradation by peptidases in the body.
XX Modifying and attaching therapeutic peptides to albumin prevents or
XX reduces the action of peptidases to increase length of activity (half
XX life) and specificity as bonding to large molecules decreases
XX intracellular uptake and interference with physiological processes.
XX AAS90929 to AAB92441 represent peptides which can be used in the
XX exemplification of the present invention
XX
XX Sequence 20 AA;
XX
XX Query Match 99.1%; Score 109; DB 4; Length 20;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-10;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 TPDINPAWYAGRGIRPVGR 19
XX | | | | | | | | | | | | | | | | | | | | | |
XX Db 1 TPDINPAWYAGRGIRPVGR 19
XX
XX RESULT 10
XX AAB46954
XX ID AAB46954 standard; protein; 20 AA.
XX
XX AC AAB46954;
XX
XX 04-MAY-2001 (first entry)
XX
XX Peptide PrRP20 fragment.
XX
XX GPR10; UHR-1; PrRP receptor; prolactin-releasing peptide; pain;
XX central nervous system disorder; autonomic regulation; analgesic;
XX hypotensive; blood pressure.
XX
XX Unidentified.
XX
XX Key Location/Qualifiers
XX FT Modified-site 20
XX FT /note= "C-terminal amide"
XX
XX WO200109182-A1.
XX
XX 08-FEB-2001.
XX
XX 03-AUG-2000; 2000WO-FI000664.
XX
XX 03-AUG-1999; 99US-00365756.
XX
XX

PR 20-MAR-2000; 2000US-005311567.
 XX (JUVV-) JUVANTIA PHARMA LTD OY.
 XX Panula PAJ, Pertovaara A, Kalso E, Korpi E;
 XX WPI; 2001-182941/18.
 XX C-terminal fragments of prolactin-releasing peptide useful for regulating
 PT autonomic functions and in the manufacture of a medicament for regulating
 PT blood pressure.
 XX Claim 1; Page 10; 40pp; English.
 XX This invention describes a novel C-terminal fragment (I) of an isolated
 CC prolactin-releasing peptide (PrRP), referred to as PrRP20 and having a
 CC sequence (SI). The invention also describes (1) a therapeutic composition
 CC (CI) comprising (I) or a C-terminal fragment of PrRP referred to as PrRP8
 CC and comprising a sequence (S2); (2) a diagnostic method based on antisera
 CC against PrRP20 for identification of disorders involving the central
 CC nervous system, including those associated with pain or autonomic
 CC regulation, where specific antisera against the N-and/or C-terminal
 CC domains of PrRP is used to identify alterations in PrRP synthesis or
 CC levels; (3) a rat or human receptor encoded by a 1122 nucleotide sequence
 CC (S3), fully defined in the specification; (4) treating a person suffering
 CC from a disorder regulated by a receptor (II) encoded by a sequence of
 CC 1122 nucleotide sequence, fully defined in the specification, located in
 CC the central nervous system, by administering an agonist or antagonist to
 CC the receptor; and (5) treating blood pressure, by blocking of receptors
 CC of PrRP or its C-terminal fragment GlyleargProvalGlyArgPhe-NH₂ (S2).
 CC The products of the invention have analgesic and hypotensive activity.
 CC (I) is useful for regulating autonomic functions, such as increasing
 CC blood pressure. (I) is useful for treating pain, for manufacturing a
 CC medicament for regulating blood pressure, and for treating pain. Agonist
 CC and antagonist of (II) are useful for treating acute pain, inflammatory
 CC pain and neuropathic pain, for regulating autonomic functions and
 CC treating high blood pressure
 XX Sequence 20 AA;
 SQ
 Query Match 99.1%; Score 109; DB 4; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TPDINPAWYAGRGIRPVGR 19
 Db |||||
 1 TPDINPAWYAGRGIRPVGR 19
 RESULT 11
 AAG62519
 ID AAG62519 standard; peptide; 20 AA.
 AC AAG62519;
 XX 24-AUG-2001 (first entry)
 DT Bovine CRH releasing protein related peptide SEQ ID NO: 6.
 DE Cow; corticotrophin releasing hormone; CRH; G protein receptor ligand;
 XX analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;
 KW Addison's disease; adrenal gland hyperfunction; obesity.
 KW Bos taurus.
 OS WO200135984-A1.
 XX 25-MAY-2001.
 PD 17-NOV-2000; 2000WC-JP0081119.
 PF 18-NOV-1999; 99JP-00327900.
 PR 26-SEP-2000; 2000JP-00297073.

XX (TAKE) TAKEDA CHEM IND LTD.
 XX Kitada C, Matsumoto H, Hinuma S;
 XX WPI; 2001-355552/37.
 XX Use of G protein receptor ligand or peptide for controlling corticotropin
 PT releasing hormone secretion.
 XX Claim 4; Page 64; 90pp; Japanese.
 XX The present sequence describes a method of controlling the secretion of
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein
 CC receptor ligand. This can be used to control the secretion of CRH and is
 CC useful as an analgesic or for treating, preventing or ameliorating
 CC diseases associated with CRH secretion such as hyperaldosteronism,
 CC hypercortisolaemia, secondary or chronic hypoadrenocorticism, Addison's
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The
 CC present sequence is a peptide used in the exemplification of the
 CC invention
 XX Sequence 20 AA;
 SQ
 Query Match 99.1%; Score 109; DB 4; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TPDINPAWYAGRGIRPVGR 19
 Db |||||
 1 TPDINPAWYAGRGIRPVGR 19
 RESULT 12
 AAE26402
 ID AAE26402 standard; peptide; 20 AA.
 XX AAE26402;
 XX 13-DEC-2002 (first entry)
 DT Bovine PrRP-31 C-terminal peptide, PrRP-20.
 XX Wakefulness; sleep disorder; prolactin releasing peptide receptor; PrRP;
 KW GPR10; therapy; epilepsy; narcolepsy; sleepiness; sleep apnoea; insomnia;
 KW idiopathic hypersomnia; psychogenic hypersomnia; seizure; anticonvulsant;
 KW bosine.
 XX Bos taurus.
 OS US2002037533-A1.
 XX 28-MAR-2002.
 PD 17-AUG-2001; 2001US-00932161.
 PF 28-APR-2000; 2000US-00560915.
 XX (CIVE/) CIVELLI O.
 XX (LINS/) LIN S.
 XX Civelli O, Lin S;
 XX WPI; 2002-403931/43.
 XX Screening for compounds useful for promoting wakefulness or sleep, and
 XX for treating sleeping disorders, e.g. insomnia, hypersomnia or sleep
 XX apnea, comprises administering a prolactin releasing peptide agonist or
 XX antagonist.
 XX Disclosure; Page 24; 35pp; English.

CC The present invention relates to a method of screening for compounds for
 CC promoting wakefulness or sleep in a mammal. The method involves
 CC administering a prolactin releasing peptide (PrRP) receptor (GPR10)
 CC agonist or antagonist respectively and determining the ability of the
 CC compound to promote wakefulness or sleep. The compounds identified from
 CC the method are used in the therapy of epilepsy and other diseases
 CC associated with absence seizures and in promoting wakefulness and sleep
 CC in individuals having sleep disorders such as insomnia and narcolepsy.
 CC PrRP receptor agonists may be used to treat common disorders which lead
 CC to sleepiness, e.g. sleep apnoea, narcolepsy, idiopathic hypersomnia and
 CC psychogenic hypersomnia. PrRP receptor antagonists are useful for
 CC promoting sleep and for treating insomnia such as adjustment sleep
 CC disorder and psychophysiological insomnia. The present sequence is bovine
 CC PrRP-31 C-terminal peptide, PrRP-20
 CC
 CC Sequence 20 AA;
 CC

Query Match 99.1%; Score 109; DB 5; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
 Db 1 TPDINPAWYAGRGIRPVGR 19
 RESULT 13
 ABU60834
 ID ABU60834 standard; peptide; 20 AA.
 XX
 AC ABU60834;
 XX
 DT 06-MAY-2003 (first entry)
 XX
 DE Peptide production by gene recombination associated peptide #18.
 XX
 KW Peptide production; low-molecular peptide; KISS-1; GPR8 ligand;
 KW gene recombination.
 XX
 OS Bos sp.
 XX
 PN WO200292829-A1.
 XX
 PD 21-NOV-2002.
 XX
 PF 16-MAY-2002; 2002WO-JP004735.
 XX
 PR 17-MAY-2001; 2001JP-00147341.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Nishimura O, Suenaga M, Ito T, Kitada C;
 XX
 DR WPI; 2003-129302/12.
 XX
 CC Process for producing peptides e.g. KISS-1 peptide and GPR8 ligand for
 CC subsequent applications by gene recombination technique through tandem
 CC repeats to provide precursor protein with specific cleavage sites.
 CC
 CC Disclosure; Page 62; 87pp; Japanese.
 CC
 CC The invention describes a method of producing a peptide comprising the
 CC excision of the N and C-terminals of a target peptide with enzymes or
 CC chemically through the attached cleavage sites repeated by ligation in a
 CC precursor protein. The method is for producing (low-molecular) peptides
 CC e.g. KISS-1 peptide and GPR8 ligand for subsequent applications by the
 CC gene recombination technique through tandem repeats to provide a
 CC precursor protein with specific cleavage sites. With this method, peptide
 CC production can be carried out easily to provide large quantities of the
 CC required peptides. This is the amino acid sequence of a peptide
 CC associated with the peptide production method of the invention
 CC
 CC Sequence 20 AA;
 CC

Query Match 99.1%; Score 109; DB 6; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
 Db 1 TPDINPAWYAGRGIRPVGR 19
 RESULT 14
 AAW31375
 ID AAW31375 standard; peptide; 21 AA.
 XX
 AC AAW31375;
 XX
 DT 06-APR-1998 (first entry)
 XX
 DE Bovine G protein-coupled receptor ligand peptide fragment 5.
 XX
 KW G protein-coupled receptor; ligand binding; pharmaceutical; modulator;
 KW pituitary; central nervous system; pancreas; prophylactic;
 KW therapeutic agent.
 XX
 OS Bos taurus.
 XX
 PN WO9724436-A2.
 XX
 PD 10-JUL-1997.
 XX
 PF 26-DEC-1996; 96WO-JP003821.
 XX
 PR 28-DEC-1995; 95JP-00343371.
 PR 15-MAR-1996; 96JP-00059419.
 PR 12-AUG-1996; 96JP-00211805.
 PR 18-SEP-1996; 96JP-00246573.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Hinuma S, Habata Y, Kawanata Y, Hosoya M, Fujii R, Fukusumi S;
 PI Kitada C;
 XX
 DR WPI; 1997-363672/33.
 DR N-PSDB; AAV02398.
 XX
 CC Ligand peptide for G protein-coupled receptor - acts by modulating
 CC function in the central nervous system, pancreas and pituitary gland.
 CC
 CC Claim 2; Page 162; 258pp; English.
 CC
 CC This sequence represents a peptide fragment of a novel bovine pituitary
 CC derived ligand corresponding to amino acid residues 34 to 54 of the
 CC sequence in AAW31368 and is used in an assay to monitor ligand binding to
 CC the G protein-coupled receptor protein. Pharmaceutical compositions
 CC containing this ligand may be used as a pituitary function modulator, a
 CC central nervous system modulator or a pancreatic function modulator. This
 CC ligand could have specific applications as a prophylactic or therapeutic
 CC agent for dementia, depression, hyperkinetic syndrome, disturbance of
 CC consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone
 CC secretory disease, hyper- and polyphagia, hypercholesterolaemia,
 CC hyperglycaemia, hyperlipidaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia,
 CC amyotrophic lateral sclerosis, acute myocardial infarction,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis, asthma, epilepsy, infertility and/or oligogalactia. Assays
 CC can also be developed to screen compounds which are capable of altering
 CC the binding activity of the ligand thus affecting activation of the G
 CC protein-coupled receptor protein
 CC
 CC Sequence 21 AA;
 CC

Query Match 99.1%; Score 109; DB 2; Length 21;

Best Local Similarity 100.0%; Pred. No. 2.2e-10; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
| | | | | | | | | | | | | | | | | | | | | |
Db 1 TPDINPAWYAGRGIRPVGR 19
| | | | | | | | | | | | | | | | | | | | | |

RESULT 15
AAW95192
ID AAW95192 standard; peptide, 21 AA.
XX
AC AAW95192;
XX
DT 10-MAR-1999 (first entry)
XX
DE Bovine pituitary-derived ligand polypeptide fragment.
XX
KW Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
KW Creutzfeldt-Jakob disease; poisoning; schizophrenia; growth hormone;
KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
KW gene therapy; transgenic animal; bovine.
XX
OS Bos sp.
XX
XX WO9849295-A1.
XX
XX 05-NOV-1998.
XX
PF 27-APR-1998; 98WO-JP001923.
XX
XX 28-APR-1997; 97JP-00109974.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Hinuma S, Fukusumi S;
XX
XX WPI; 1999-009423/01.
XX
XX New polypeptide ligand for orphan G protein coupled receptors - used for
XX treating disorders of central nervous system, pituitary and pancreas, and
XX for drug screening.
XX
XX Example 19; Page 151; 206pp; English.
XX
XX The invention relates to a murine pituitary-derived ligand polypeptide
XX which is a ligand for the G-protein coupled orphan receptor designated
XX GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing
XX the ligand polypeptide encoding DNA are used to produce a recombinant
XX ligand polypeptide. The ligand polypeptide, and its fragments, modulate
XX function of the pituitary, central nervous system, pancreas and other
XX tissues and can be used to screen for agents that modulate binding of the
XX polypeptide to the receptor; to quantify the amount of receptor in a
XX sample and to raise antibodies. They may also be used therapeutically,
XX e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's
XX diseases; Creutzfeldt-Jakob disease; poisoning by heavy metals or drugs;
XX diabetes; schizophrenia; disorders of growth hormone secretion; cancer;
XX rheumatoid arthritis, epilepsy and many others, also to improve post-
XX operative nutritional status and as vasopressor. Transgenic animals
XX carrying the ligand polypeptide encoding DNA or its mutein are used to
XX study the function of the polypeptide-expressing genes, as models of
XX disease, for drug screening and as source of cell lines. The ligand
XX polypeptide DNA is used as a source of probes and primers; to identify
XX related sequences; in receptor-binding assays; for production of Ab and
XX antisera; in drug development; for gene therapy and to develop transgenic
XX animals. The present sequence represents a bovine genome-derived ligand
XX polypeptide fragment which is similar to the murine ligand-polypeptide
XX Sequence 21 AA;

Query Match 99.1%; Score 109; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.2e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
| | | | | | | | | | | | | | | | | | | | | |
Db 1 TPDINPAWYAGRGIRPVGR 19
| | | | | | | | | | | | | | | | | | | | | |

Search completed: August 12, 2004, 14:43:55
Job time : 42.5669 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:49:10 ; Search time 35.3488 Seconds
(without alignments)
177.617 Million cell updates/sec

Title: US-09-700-643A-12

Perfect score: 110

Sequence: 1 TPDINPAWYAGRGIRPVGR 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1292805 seqs, 313927144 residues

Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCT05_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109	99.1	19	13	US-10-044-592-27
2	109	99.1	20	9	US-09-932-161-16
3	109	99.1	20	13	US-10-044-592-42
4	109	99.1	20	14	US-10-096-777-16
5	109	99.1	21	13	US-10-044-592-43
6	109	99.1	22	13	US-10-044-592-44
7	109	99.1	31	9	US-09-932-161-13
8	109	99.1	31	13	US-10-044-592-39
9	109	99.1	31	14	US-10-096-777-13
10	109	99.1	32	13	US-10-044-592-40
11	109	99.1	32	13	US-10-044-592-41
12	109	99.1	98	13	US-10-044-592-28
13	109	99.1	98	13	US-10-044-592-38
14	109	99.1	98	13	US-10-044-592-82
15	109	99.1	98	13	US-10-044-592-84

16	109	99.1	98	13	US-10-044-592-86	Sequence 86, Appl
17	109	99.1	98	13	US-10-044-592-88	Sequence 88, Appl
18	105	95.5	20	9	US-09-932-161-17	Sequence 17, Appl
19	105	95.5	20	13	US-10-044-592-6	Sequence 6, Appl
20	105	95.5	20	14	US-10-096-777-17	Sequence 17, Appl
21	105	95.5	31	9	US-09-932-161-14	Sequence 14, Appl
22	105	95.5	31	13	US-10-044-592-4	Sequence 4, Appl
23	105	95.5	31	13	US-10-044-592-5	Sequence 5, Appl
24	105	95.5	31	14	US-10-096-777-14	Sequence 14, Appl
25	105	95.5	70	13	US-10-044-592-90	Sequence 90, Appl
26	105	95.5	82	13	US-10-044-592-1	Sequence 1, Appl
27	105	95.5	86	13	US-10-044-592-96	Sequence 96, Appl
28	105	95.5	91	13	US-10-044-592-94	Sequence 94, Appl
29	104	94.5	23	13	US-10-044-592-26	Sequence 26, Appl
30	103	93.6	20	9	US-09-932-161-18	Sequence 18, Appl
31	103	93.6	20	14	US-10-096-777-18	Sequence 18, Appl
32	103	93.6	31	9	US-09-932-161-15	Sequence 15, Appl
33	103	93.6	31	14	US-10-096-777-15	Sequence 92, Appl
34	103	93.6	87	13	US-10-044-592-92	Sequence 78, Appl
35	94	85.5	25	13	US-10-044-592-78	Sequence 8, Appl
36	57	51.8	9	13	US-10-044-592-8	Sequence 197, Appl
37	52	47.3	465	14	US-10-301-822-197	Sequence 268092,
38	52	47.3	183	12	US-10-424-599-268092	Sequence 4642, Ap
39	50	45.5	417	9	US-09-738-626-4642	Sequence 7594, Ap
40	50	45.5	130	14	US-10-156-761-7594	Sequence 6327, Ap
41	50	45.5	527	15	US-10-369-493-6327	Sequence 469, App
42	48.5	44.1	664	12	US-10-389-647-469	Sequence 9, Appl
43	48	43.6	10	13	US-10-044-592-9	Sequence 10748, A
44	48	43.6	402	14	US-10-156-761-10748	Sequence 151465,
45	47	42.7	140	12	US-10-424-599-151465	

ALIGNMENTS

RESULT 1
US-10-044-592-27
; Sequence 27, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JF98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 27
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-27

Query Match 99.1%; Score 109; DB 13; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
Db 1 TPDINPAWYAGRGIRPVGR 19

RESULT 2
US-09-932-161-16
; Sequence 16, Application US/09932161
; Patent No. US20020037533A1

```

; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-932-161-16

Query Match          99.1%; Score 109; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRGIRPVGR 19
Db 1 TPDINPAWYAGRGIRPVGR 19

RESULT 3
US-10-044-592-42
; Sequence 42, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Himuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 42
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-42

Query Match          99.1%; Score 109; DB 13; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRGIRPVGR 19
Db 1 TPDINPAWYAGRGIRPVGR 19

RESULT 4
US-10-096-777-16
; Sequence 16, Application US/10096777
; Publication No. US20030171270A1
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777
; CURRENT FILING DATE: 2002-03-12

```

```

; PRIOR APPLICATION NUMBER: US/09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Bos taurus
US-10-096-777-16

```

```

Query Match          99.1%; Score 109; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 TPDINPAWYAGRGIRPVGR 19
Db 1 TPDINPAWYAGRGIRPVGR 19

```

```

RESULT 5
US-10-044-592-43
; Sequence 43, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Himuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 43
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-43

```

```

Query Match          99.1%; Score 109; DB 13; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 TPDINPAWYAGRGIRPVGR 19
Db 1 TPDINPAWYAGRGIRPVGR 19

```

```

RESULT 6
US-10-044-592-44
; Sequence 44, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Himuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 44

```

; LENGTH: 22
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-44

Query Match 99.1%; Score 109; DB 13; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.1e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
| | | | | | | | | | | | | | | | | | | | | |
DB 1 TPDINPAWYAGRGIRPVGR 19

RESULT 7

US-09-932-161-13
; Sequence 13, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Civeilli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; TITLE OF INVENTION: Promoting Wakefulness and Sleep
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-932-161-13

Query Match 99.1%; Score 109; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. No. 3e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
| | | | | | | | | | | | | | | | | | | | | |
DB 12 TPDINPAWYAGRGIRPVGR 30

RESULT 8

US-10-044-592-39
; Sequence 39, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 39
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-39

Query Match 99.1%; Score 109; DB 13; Length 31;
Best Local Similarity 100.0%; Pred. No. 3e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
| | | | | | | | | | | | | | | | | | | | | |
DB 12 TPDINPAWYAGRGIRPVGR 30

RESULT 9

US-10-096-777-13
; Sequence 13, Application US/10096777
; Publication No. US20030171270A1
; GENERAL INFORMATION:
; APPLICANT: Civeilli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; TITLE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Bos taurus
US-10-096-777-13

Query Match 99.1%; Score 109; DB 14; Length 31;
Best Local Similarity 100.0%; Pred. No. 3e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
| | | | | | | | | | | | | | | | | | | | | |
DB 12 TPDINPAWYAGRGIRPVGR 30

RESULT 10

US-10-044-592-40
; Sequence 40, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 40
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-40

Query Match 99.1%; Score 109; DB 13; Length 32;
Best Local Similarity 100.0%; Pred. No. 3.1e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
| | | | | | | | | | | | | | | | | | | | | |
DB 12 TPDINPAWYAGRGIRPVGR 30

RESULT 11

US-10-044-592-41

; Sequence 41, Application US/10044592
 ; Publication No. US20020143152A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hinuma, Shuji
 ; TITLE OF INVENTION: Polypeptides, their Production and Use
 ; FILE REFERENCE: 2463US2P
 ; CURRENT APPLICATION NUMBER: US/10/044,592
 ; CURRENT FILING DATE: 2002-01-10
 ; PRIOR FILING DATE: 1999-25-10
 ; PRIOR APPLICATION NUMBER: US 09/403639
 ; PRIOR FILING DATE: 1998-04-27
 ; PRIOR APPLICATION NUMBER: JP 9-109974
 ; PRIOR FILING DATE: 1997-04-28
 ; NUMBER OF SEQ ID NOS: 96
 ; SOFTWARE:
 ; SEQ ID NO 41
 ; LENGTH: 33
 ; TYPE: PRT
 ; ORGANISM: Bovine
 US-10-044-592-41

Query Match 99.1%; Score 109; DB 13; Length 33;
 Best Local Similarity 100.0%; Pred. No. 3.2e-09;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
 |||||
 Db 12 TPDINPAWYAGRGIRPVGR 30

RESULT 12

US-10-044-592-28
 ; Sequence 28, Application US/10044592
 ; Publication No. US20020143152A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hinuma, Shuji
 ; TITLE OF INVENTION: Polypeptides, their Production and Use
 ; FILE REFERENCE: 2463US2P
 ; CURRENT APPLICATION NUMBER: US/10/044,592
 ; CURRENT FILING DATE: 2002-01-10
 ; PRIOR FILING DATE: 1999-25-10
 ; PRIOR APPLICATION NUMBER: US 09/403639
 ; PRIOR FILING DATE: 1998-04-27
 ; PRIOR APPLICATION NUMBER: JP 9-109974
 ; PRIOR FILING DATE: 1997-04-28
 ; NUMBER OF SEQ ID NOS: 96
 ; SOFTWARE:
 ; SEQ ID NO 28
 ; LENGTH: 98
 ; TYPE: PRT
 ; ORGANISM: Murine
 US-10-044-592-28

Query Match 99.1%; Score 109; DB 13; Length 98;
 Best Local Similarity 100.0%; Pred. No. 9.3e-09;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
 |||||
 Db 34 TPDINPAWYAGRGIRPVGR 52

RESULT 13

US-10-044-592-38
 ; Sequence 38, Application US/10044592
 ; Publication No. US20020143152A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hinuma, Shuji
 ; TITLE OF INVENTION: Polypeptides, their Production and Use
 ; FILE REFERENCE: 2463US2P
 ; CURRENT APPLICATION NUMBER: US/10/044,592
 ; CURRENT FILING DATE: 2002-01-10
 ; PRIOR APPLICATION NUMBER: US 09/403639
 ; PRIOR FILING DATE: 1997-04-28
 ; NUMBER OF SEQ ID NOS: 96
 ; SOFTWARE:
 ; SEQ ID NO 38
 ; LENGTH: 98
 ; TYPE: PRT
 ; ORGANISM: Murine
 US-10-044-592-38

; TITLE OF INVENTION: Polypeptides, their Production and Use
 ; FILE REFERENCE: 2463US2P
 ; CURRENT APPLICATION NUMBER: US/10/044,592
 ; CURRENT FILING DATE: 2002-01-10
 ; PRIOR FILING DATE: 1999-25-10
 ; PRIOR APPLICATION NUMBER: US 09/403639
 ; PRIOR FILING DATE: 1998-04-27
 ; PRIOR APPLICATION NUMBER: JP 9-109974
 ; PRIOR FILING DATE: 1997-04-28
 ; NUMBER OF SEQ ID NOS: 96
 ; SOFTWARE:
 ; SEQ ID NO 38
 ; LENGTH: 98
 ; TYPE: PRT
 ; ORGANISM: Bovine
 US-10-044-592-38

Query Match 99.1%; Score 109; DB 13; Length 98;
 Best Local Similarity 100.0%; Pred. No. 9.3e-09;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
 |||||
 Db 34 TPDINPAWYAGRGIRPVGR 52

RESULT 14

US-10-044-592-82
 ; Sequence 82, Application US/10044592
 ; Publication No. US20020143152A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hinuma, Shuji
 ; TITLE OF INVENTION: Polypeptides, their Production and Use
 ; FILE REFERENCE: 2463US2P
 ; CURRENT APPLICATION NUMBER: US/10/044,592
 ; CURRENT FILING DATE: 2002-01-10
 ; PRIOR FILING DATE: 1999-25-10
 ; PRIOR APPLICATION NUMBER: US 09/403639
 ; PRIOR FILING DATE: 1998-04-27
 ; PRIOR APPLICATION NUMBER: JP 9-109974
 ; PRIOR FILING DATE: 1997-04-28
 ; NUMBER OF SEQ ID NOS: 96
 ; SOFTWARE:
 ; SEQ ID NO 82
 ; LENGTH: 98
 ; TYPE: PRT
 ; ORGANISM: Bovine
 US-10-044-592-82

Query Match 99.1%; Score 109; DB 13; Length 98;
 Best Local Similarity 100.0%; Pred. No. 9.3e-09;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
 |||||
 Db 34 TPDINPAWYAGRGIRPVGR 52

RESULT 15

US-10-044-592-84
 ; Sequence 84, Application US/10044592
 ; Publication No. US20020143152A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hinuma, Shuji
 ; TITLE OF INVENTION: Polypeptides, their Production and Use
 ; FILE REFERENCE: 2463US2P
 ; CURRENT APPLICATION NUMBER: US/10/044,592
 ; CURRENT FILING DATE: 2002-01-10
 ; PRIOR APPLICATION NUMBER: US 09/403639
 ; PRIOR FILING DATE: 1997-04-28
 ; NUMBER OF SEQ ID NOS: 96
 ; SOFTWARE:
 ; SEQ ID NO 84
 ; LENGTH: 98
 ; TYPE: PRT
 ; ORGANISM: Bovine
 US-10-044-592-84

Query Match 99.1%; Score 109; DB 13; Length 98;
 Best Local Similarity 100.0%; Pred. No. 9.3e-09;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
 |||||
 Db 34 TPDINPAWYAGRGIRPVGR 52

```

; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 84
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-84

Query Match      99.1%; Score 109; DB 13; Length 98;
Best Local Similarity 100.0%; Pred. No. 9.3e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TPDINPAWYAGRGIRPVGR 19
Db      34 TPDINPAWYAGRGIRPVGR 52

Search completed: August 12, 2004, 15:22:51
Job time : 35.3488 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:36 ; Search time 11.3953 Seconds
(without alignments)
90.609 Million cell updates/sec

Title: US-09-700-643A-12

Perfect score: 110

Sequence: 1 TPDINPAWYAGRGIRPVGRX 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A COMB pep:*
2: /cgn2_6/prodata/2/iaa/5B COMB pep:*
3: /cgn2_6/prodata/2/iaa/6A COMB pep:*
4: /cgn2_6/prodata/2/iaa/6B COMB pep:*
5: /cgn2_6/prodata/2/iaa/PTUS COMB pep:*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	109	99.1	19	3	US-09-105-678A-30
2	109	99.1	19	3	US-08-776-971-4
3	109	99.1	19	3	US-09-421-208-30
4	109	99.1	20	3	US-09-105-678A-34
5	109	99.1	20	3	US-08-776-971-8
6	109	99.1	20	3	US-08-776-971-98
7	109	99.1	20	3	US-09-421-208-34
8	109	99.1	20	4	US-09-560-915-16
9	109	99.1	21	3	US-09-105-678A-35
10	109	99.1	21	3	US-08-776-971-9
11	109	99.1	21	3	US-09-421-208-35
12	109	99.1	22	3	US-09-105-678A-36
13	109	99.1	22	3	US-08-776-971-10
14	109	99.1	22	3	US-09-421-208-36
15	109	99.1	31	3	US-09-105-678A-7
16	109	99.1	31	3	US-09-105-678A-31
17	109	99.1	31	3	US-08-776-971-5
18	109	99.1	31	3	US-08-776-971-97
19	109	99.1	31	3	US-09-421-208-7
20	109	99.1	31	3	US-09-421-208-31
21	109	99.1	31	4	US-09-560-915-13
22	109	99.1	32	3	US-09-105-678A-32
23	109	99.1	32	3	US-08-776-971-6
24	109	99.1	33	3	US-09-421-208-32
25	109	99.1	33	3	US-09-105-678A-33
26	109	99.1	33	3	US-08-776-971-7
27	109	99.1	33	3	US-09-421-208-33

Sequence 1, Appli
Sequence 44, Appl
Sequence 122, App
Sequence 131, App
Sequence 136, App
Sequence 40, Appl
Sequence 50, Appl
Sequence 40, Appl
Sequence 17, Appl
Sequence 41, Appl
Sequence 51, Appl
Sequence 42, Appl
Sequence 52, Appl
Sequence 42, Appl
Sequence 8, Appl
Sequence 37, Appl
Sequence 4, Appl

28 109 99.1 98 3 US-08-776-971-1
29 109 99.1 98 3 US-08-776-971-44
30 109 99.1 98 3 US-08-776-971-122
31 109 99.1 98 3 US-08-776-971-131
32 109 99.1 98 3 US-08-776-971-136
33 105 95.5 20 3 US-09-105-678A-40
34 105 95.5 20 3 US-08-776-971-50
35 105 95.5 20 3 US-09-421-208-40
36 105 95.5 20 4 US-09-560-915-17
37 105 95.5 21 3 US-09-105-678A-41
38 105 95.5 21 3 US-08-776-971-51
39 105 95.5 21 3 US-09-421-208-41
40 105 95.5 22 3 US-09-105-678A-42
41 105 95.5 22 3 US-08-776-971-52
42 105 95.5 22 3 US-09-421-208-42
43 105 95.5 31 3 US-09-105-678A-8
44 105 95.5 31 3 US-09-105-678A-37
45 105 95.5 31 3 US-09-172-353-4

ALIGNMENTS

RESULT 1
US-09-105-678A-30
; Sequence 30, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/105.678A
; APPLICATION NUMBER: US/09/105.678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-105-678A-30

Query Match 99.1%; Score 109; DB 3; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.8e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19

Db 1 TPDINPAWYAGRGIRPVGR 19
|||||
RESULT 2
US-08-776-971-4
; Sequence 4, Application US/08776971B
; Patent No. 622884
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; Habata, Yugo
; Kawamata, Yuji
; Hosoya, Masaki
; Fujii, Ryo
; Fukusumi, Shoji
; Kitada, Chieko
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,971B
FILING DATE: 06-Feb-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03821
FILING DATE: 28-DEC-1996
APPLICATION NUMBER: JP 7/343371
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: JP 8/59419
FILING DATE: 15-MAR-1996
APPLICATION NUMBER: JP 8/211805
FILING DATE: 12-AUG-1996
APPLICATION NUMBER: JP 8/246573
FILING DATE: 18-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 47176
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-776-971-4
Query Match 99.1%; Score 109; DB 3; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.8e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TPDINPAWYAGRGIRPVGR 19
Db 1 TPDINPAWYAGRGIRPVGR 19
|||||
RESULT 3
US-08-776-971-4
; Sequence 4, Application US/08776971B
; Patent No. 622884
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; Habata, Yugo
; Kawamata, Yuji
; Hosoya, Masaki
; Fujii, Ryo
; Fukusumi, Shoji
; Kitada, Chieko
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,971B
FILING DATE: 06-Feb-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03821
FILING DATE: 28-DEC-1996
APPLICATION NUMBER: JP 7/343371
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: JP 8/59419
FILING DATE: 15-MAR-1996
APPLICATION NUMBER: JP 8/211805
FILING DATE: 12-AUG-1996
APPLICATION NUMBER: JP 8/246573
FILING DATE: 18-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 47176
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-776-971-4
Query Match 99.1%; Score 109; DB 3; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.8e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TPDINPAWYAGRGIRPVGR 19
Db 1 TPDINPAWYAGRGIRPVGR 19
|||||
RESULT 4
US-09-105-678A-34
; Sequence 34, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

US-09-421-208-30
; Sequence 30, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/421,208
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/105,678
FILING DATE: 26-JUN-1998
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-421-208-30
Query Match 99.1%; Score 109; DB 3; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.8e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TPDINPAWYAGRGIRPVGR 19
Db 1 TPDINPAWYAGRGIRPVGR 19
|||||
RESULT 4
US-09-105-678A-34
; Sequence 34, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: linear
MOLECULE TYPE: peptide
US-09-105-678A-34

Query Match 99.1%; Score 109; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.2e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRGIRPVGR 19
Db 1 TPDINPAWYAGRGIRPVGR 19

RESULT 5
US-08-776-971-8
Sequence 8, Application US/08776971B
Patent No. 6228984
GENERAL INFORMATION:
APPLICANT: Hinuma, Shuji
Habata, Yugo
Kawamata, Yuji
Hosoya, Masaki
Fujii, Ryo
Fukusumi, Shoji
Kitada, Chieko
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,971B
FILING DATE: 06-Feb-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03821
FILING DATE: 28-DEC-1996
APPLICATION NUMBER: JP 7/343371
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: JP 8/59419
FILING DATE: 15-MAR-1996
APPLICATION NUMBER: JP 8/211805
FILING DATE: 12-AUG-1996
APPLICATION NUMBER: JP 8/246573
FILING DATE: 18-SEP-1996
ATTORNEY/AGENT INFORMATION:

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: linear
MOLECULE TYPE: peptide
US-09-105-678A-34

Query Match 99.1%; Score 109; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.2e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRGIRPVGR 19
Db 1 TPDINPAWYAGRGIRPVGR 19

RESULT 5
US-08-776-971-8
Sequence 8, Application US/08776971B
Patent No. 6228984
GENERAL INFORMATION:
APPLICANT: Hinuma, Shuji
Habata, Yugo
Kawamata, Yuji
Hosoya, Masaki
Fujii, Ryo
Fukusumi, Shoji
Kitada, Chieko
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,971B
FILING DATE: 06-Feb-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03821
FILING DATE: 28-DEC-1996
APPLICATION NUMBER: JP 7/343371
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: JP 8/59419
FILING DATE: 15-MAR-1996
APPLICATION NUMBER: JP 8/211805
FILING DATE: 12-AUG-1996
APPLICATION NUMBER: JP 8/246573
FILING DATE: 18-SEP-1996
ATTORNEY/AGENT INFORMATION:

```

, NAME: Conlin, David G.
, REGISTRATION NUMBER: 27,026
, REFERENCE/DOCKET NUMBER: 47176
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: 617-523-3400
, TELEFAX: 617-523-6440
, INFORMATION FOR SEQ ID NO: 98:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 20 amino acids
, TYPE: amino acid
, STRANDEDNESS: single
, TOPOLOGY: linear
, SEQUENCE DESCRIPTION: SEQ ID NO: 98:
US-08-776-971-98

```

Query Match 99.1%; Score 109; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.2e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRGIRPVGR 19
|||
Db 1 TPDINPAWYAGRGIRPVGR 19

RESULT 7

US-09-421-208-34 ;
; Sequence 34, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street

CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/421,208
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 09/105,678
FILING DATE: 26-JUN-1998
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-421-208-34

	Matches	19;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	TPDINPAWYAGRGIRPVGR	19							
Db	1	TPDINPAWYAGRGIRPVGR	19							

RESULT 8

```

US-09-560-915-16
; Sequence 16, Application US/09560915
; Patent No. 6383764
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; TITLE OF INVENTION: Relating To Prolactin Releasing Peptide (Prp)
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/09/560,915
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-560-915-16

```

```
Query Match      99.1%; Score 109; DB 4; Length 20;
Best Local Similarity 100.0%; Pred.No. 7.2e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 TPDINPAWYAGRGIRPVGR 19
|||
Db 1 TPDINPAWYAGRGIRPVGR 19

RESULT. T 9

US-09-105-678A-35
; Sequence 35, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takao
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 172119/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 35:

```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-105-678A-35

Query Match 99.1%; Score 109; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.6e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRGIRPVGR 19
Db 1 TPDINPAWYAGRGIRPVGR 19

RESULT 10
US-08-776-971-9
; Sequence 9, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; HABATA, Yugo
; KAWAMATA, YUJI
; Hosoya, Masaki
; Fujii, Ryo
; Fukusumi, Shoji
; Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
;

; SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-08-776-971-9
Query Match 99.1%; Score 109; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.6e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRGIRPVGR 19
Db 1 TPDINPAWYAGRGIRPVGR 19

RESULT 11
US-09-421-208-35
; Sequence 35, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-421-208-35

Query Match 99.1%; Score 109; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.6e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRGIRPVGR 19
Db 1 TPDINPAWYAGRGIRPVGR 19

RESULT 12
US-09-105-678A-36
; Sequence 36, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
;
```

APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-36

Query Match 99.1%; Score 109; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 8e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
Db 1 TPDINPAWYAGRGIRPVGR 19

RESULT 13
US-08-776-971-10
Sequence 10, Application US/08776971B
Patent No. 6228984
GENERAL INFORMATION:
APPLICANT: Hinuma, Shuji
Habata, Yugo
Kawamata, Yuji
Hosoya, Masaki
Fujii, Ryo
Fukushima, Shoichi
Kitada, Chieko
TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible

OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,971B
FILING DATE: 06-Feb-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03821
FILING DATE: 28-DEC-1996
APPLICATION NUMBER: JP 7/343371
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: JP 8/59419
FILING DATE: 15-MAR-1996
APPLICATION NUMBER: JP 8/211805
FILING DATE: 12-AUG-1996
APPLICATION NUMBER: JP 8/246573
FILING DATE: 18-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 47176
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-08-776-971-10

Query Match 99.1%; Score 109; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 8e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
Db 1 TPDINPAWYAGRGIRPVGR 19

RESULT 14
US-09-421-208-36
Sequence 36, Application US/09421208
Patent No. 6258561
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/421,208
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/105,678
FILING DATE: 26-JUN-1998

Thu Aug 12 15:25:11 2004

APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-421-208-36

Query Match 99.1%; Score 109; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 8e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
| | | | | | | | | | | | | | | | | | | | |
Db 1 TPDINPAWYAGRGIRPVGR 19

RESULT 15
US-09-105-678A-7
Sequence 7, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSES: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-7

Query Match 99.1%; Score 109; DB 3; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.2e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGR 19
| | | | | | | | | | | | | | | | | | | | |
Db 12 TPDINPAWYAGRGIRPVGR 30

Search completed: August 12, 2004, 14:52:14
Job time : 11.3953 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:36 ; Search time 5.74419 Seconds
(without alignments)
217.697 Million cell updates/sec

Title: US-09-700-643A-1_COPY_12_24

Perfect score: 78

Sequence: 1 TPDINPAWYAGRG 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR 78:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	94.9	83	JC7607	prolactin-releasin
2	50	64.1	527	T33175	hypothetical prote
3	47	60.3	430	B69009	conserved hypothet
4	45	57.7	767	T21969	hypothetical prote
5	44	56.4	398	C84780	hypothetical prote
6	43	55.1	419	AH3166	hypothetical prote
7	43	55.1	511	B37222	cytochrome P450 1A
8	43	55.1	548	T47548	hypothetical prote
9	42	53.8	311	S66600	cytochrome-c oxida
10	42	53.8	335	E87151	LytB homolog [impo
11	42	53.8	375	F91173	probable transport
12	42	53.8	375	F86019	probable transport
13	42	53.8	375	S47704	hypothetical 41.1K
14	41.5	53.2	779	T49717	related to BCS1 pr
15	41	52.6	126	B72621	hypothetical prote
16	41	52.6	139	H84384	hypothetical prote
17	41	52.6	313	S57669	proclavaminic acid
18	41	52.6	391	I50702	transcription fact
19	41	52.6	395	AE3349	hypothetical prote
20	41	52.6	476	G64720	probable amino aci
21	41	52.6	476	G90629	probable inner mem
22	41	52.6	476	G85480	inner membrane tra
23	41	52.6	529	T51214	related to trichod
24	41	52.6	579	E83144	hypothetical prote
25	41	52.6	986	1 OYRGA	sparact receptor p
26	41	52.6	1425	T30811	hepatocyte growth
27	40.5	51.9	664	F83376	conserved hypothet
28	40	51.3	120	T51754	endo-xyloglucan tr
29	40	51.3	246	G75570	conserved hypothet

ALIGNMENTS

RESULT 1

JC7607

prolactin-releasing peptide - rat

C/Species: Rattus norvegicus (Norway rat)

C/Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001

C/Accession: JC7607

R/Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Hosoya, T

Biochem. Biophys. Res. Commun. 281, 53-56, 2001

A/Title: Isolation and characterization of the rat prolactin-releasing peptide gene: Mul

A/Reference number: JC7607; MUID:21092785; PMID:11178959

A/Contents: Spleen

A/Accession: JC7607

A/Molecule type: DNA

A/Residues: 1-83 <YAM>

A/Cross-references: DDBJ:AB040612; DDBJ:AB040613

C/Comment: This peptide induces arachidonic acid metabolite release from rat anterior pituitary.

C/Genetics:

A/Gene: PrRP

A/Introns: 33/1

Query Match 94.9%; Score 74; DB 2; Length 83;
Best Local Similarity 92.3%; Pred. No. 5e-05; Mismatches 0; Indels 1; Gaps 0;
Matches 12; Conservative 0

Qy 1 TPDINPAWYAGRG 13

Db 33 TPDINPAWYTGGR 45

RESULT 2

T33175

hypothetical protein C24G6.6 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C/Accession: T33175

R/Greco, T.; Bradshaw, H.; Keppler, D.

submitted to the EMBL Data Library, May 1998

A/Description: The sequence of C. elegans cosmid C24G6.

A/Reference number: Z21298

A/Accession: T33175

A/Status: preliminary; translated from GB/EMBL/DDBJ

A/Molecule type: DNA

A/Residues: 1-527 <GRE>

A/Cross-references: EMBL:AF067936; PDB:AAI9213.1; GSFDB:GNO0023; CESP:C24G6.6

A/Experimental source: strain Bristol N2; clone C24G6

C/Genetics:

A/Gene: CESP:C24G6.6

A/Map position: 5

A/Introns: 20/3; 77/1; 129/2; 208/3; 470/2

Query Match 64.1%; Score 50; DB 2; Length 527;

```

Best Local Similarity 66.7%; Pred. No. 2.4;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 PDINPAWYAGR 13
Db 370 PNLVSAWYAGR 381

RESULT 3
B69009
conserved hypothetical protein MTH1070 - Methanobacterium thermoautotrophicum (strain De
C;Species: Methanobacterium thermoautotrophicum
C;Date: 28-Jan-1999 #sequence_revision 29-Jan-1999 #text_change 23-Jul-1999
C;Accession: B69009
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: B69009
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-430 <MTH>
A;Cross-references: GB:AE000877; GB:AE000666; NID:G2622157; PIDN:AA885559.1; PID:G262216
A;Experimental source: strain Delta H
C;Genetics:
C;Superfamily: conserved hypothetical protein MTH1070

Query Match 60.3%; Score 47; DB 1; Length 430;
Best Local Similarity 80.0%; Pred. No. 6.1;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 DINPAWYAGR 12
Db 191 DINPEWAGR 200

RESULT 4
T21969
hypothetical protein F38E11.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T21969
R;Matthews, P.
Submitted to the EMBL Data Library, January 1996
A;Reference number: Z19495
A;Accession: T21969
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-767 <WIL>
A;Cross-references: EMBL:Z68342; PIDN:CAA92775.1; GSPDB:GN00022; CESP:F38E11.7
A;Experimental source: clone F38E11
C;Genetics:
A;Gene: CESP:F38E11.7
A;Map position: 4
A;Introns: 50/2; 118/1; 139/2; 189/3; 226/1; 248/1; 287/2; 375/2; 432/3; 465/3; 548/1; 6

Query Match 57.7%; Score 45; DB 2; Length 767;
Best Local Similarity 75.0%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PDINPAWY 9
Db 747 PDVKEAWY 754

RESULT 5
C84780
hypothetical protein At2g36400 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

```

```

C;Accession: C84780
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: C84780
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-398 <STO>
A;Cross-references: GB:AE002093; NID:G4581140; PIDN:AA24624.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g36400
A;Map position: 2

Query Match 56.4%; Score 44; DB 2; Length 398;
Best Local Similarity 53.6%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 PDINPAWYAGR 12
Db 130 PHYQPAWYAGR 140

RESULT 6
AH3166
hypothetical protein ugpB [imported] - Agrobacterium tumefaciens (strain C58, Dupont) pl
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AH3166
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Giller, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AH3166
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-419 <KUR>
A;Cross-references: GB:AE008687; PIDN:AAL45750.1; PID:G17743483; GSPDB:GN00188
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: ugpB
A;Genome: plasmid

Query Match 55.1%; Score 43; DB 2; Length 419;
Best Local Similarity 46.2%; Pred. No. 27;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGR 13
Db 323 SPEANATWFAAGT 335

RESULT 7
B37222
cytochrome P450 1A2, hepatic - dog
N;Alternate names: cytochrome P450 (Dah2); cytochrome P450-D3
N;Contains: oxidoreductase (EC 1.-.-)
C;Species: Canis lupus familiaris (dog)
C;Date: 31-Mar-1992 #sequence_revision 01-Mar-1996 #text_change 04-Mar-2000
C;Accession: B37222; A60463
R;Uchida, T.; Komori, M.; Kitada, M.; Kamataki, T.
Mol. Pharmacol. 38, 644-651, 1990
A;Title: Isolation of cDNAs coding for three different forms of liver microsomal cytochr
A;Reference number: A37222; MUID:91042464; PMID:2122230
A;Accession: B37222
A;Status: not compared with conceptual translation
A;Molecule type: mRNA

```

A:Residues: 10-511 <UCH>
 R:Ohta, K.; Motoya, M.; Komori, M.; Miura, T.; Kitada, M.; Kamataki, T.
 Biochem. Pharmacol. 38, 91-96, 1989
 A:Title: A novel form of cytochrome P-450 in beagle dogs. P-450-D3 is a low spin form of
 A:Reference number: A60463; MUID:89087526; PMID:2910310
 A:Accession: A60463
 A:Molecule type: protein
 A:Residues: 1-13, 'A', 15-16 <OHT>
 C:Genetics:
 A:Gene: CYP1A2
 C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
 C:Keywords: chromoprotein; electron transfer; heme; iron; liver; metalloprotein; monooxy
 F:309-475/Domain: cytochrome P450 homology <P45>
 F:453/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 55.1%; Score 43; DB 2; Length 511;
 Best Local Similarity 58.3%; Pred. No. 32;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TPDPNPANYAGR 12
 :|||
 DB 121 SPDSGFVWAAGR 132
 :|||

RESULT 8
 T47548
 hypothetical protein F8J2.80 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T47548
 R:Nyakatura, G.; Fartmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichselgartner, M.;
 Mayer, K.P.X.
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: Z4458
 A:Accession: T47548
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-548 <NYA>
 A:Cross-references: EMBL:AL132969
 A:Experimental source: cultivar Columbia; BAC clone F8J2
 C:Genetics:
 A:Map position: 3
 A:Introns: 78/1; 143/1; 242/2; 377/3; 428/2; 447/3; 470/3; 502/3; 532/3
 A:Note: F8J2.80

Query Match 55.1%; Score 43; DB 2; Length 548;
 Best Local Similarity 58.3%; Pred. No. 35;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 PDINPAWYAGR 13
 :|||
 DB 137 PHQPSWYWGRG 148
 :|||

RESULT 9
 S66600
 cytochrome-c oxidase (EC 1.9.3.1) chain III - blue mussel mitochondrion
 C:Species: mitochondrion Mytilus edulis (blue mussel)
 C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999
 C:Accession: S66600; S28753
 R:Hoffmann, R.J.; Moore, J.L.; Brown, W.M.
 submitted to the EMBL Data Library, June 1992
 A:Reference number: S66600
 A:Accession: S66600
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-311 <HOF>
 A:Cross-references: EMBL:M83760; NID:9342485; PIDN:AA31907.1; PID:G342488
 R:Hoffmann, R.J.; Moore, J.L.; Brown, W.M.
 Genetics 131, 397-412, 1992
 A:Title: A novel mitochondrial genome organization for the blue mussel, Mytilus edulis.
 A:Reference number: S28743; MUID:92354892; PMID:1386586
 A:Accession: S28753

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-163, 'N', 165-311 <BRO>
 A:Cross-references: EMBL:M83760
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC4
 C:Superfamily: cytochrome-c oxidase chain III
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane

Query Match 53.8%; Score 42; DB 2; Length 311;
 Best Local Similarity 66.7%; Pred. No. 29;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PDINPAWYA 10
 :|||
 DB 284 PDAKPSWYA 292
 :|||

RESULT 10
 E87151
 LytB homolog [imported] - Mycobacterium leprae
 C:Species: Mycobacterium leprae
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
 C:Accession: E87151
 R: Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
 sam, M.A.; Rutherford, K.M.
 Nature 409, 1007-1011, 2001
 A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Su
 A:Title: Massive gene decay in the leprosy bacillus.
 A:Reference number: A86909; MUID:21128732; PMID:11234002
 A:Accession: E87151
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-335 <STO>
 A:Cross-references: GB:AL450380; NID:gl3093596; PIDN:CAC30893.1; GSPDB:GN00147
 C:Genetics:
 A:Gene: lytB2
 C:Superfamily: penicillin tolerance protein

Query Match 53.8%; Score 42; DB 2; Length 335;
 Best Local Similarity 77.8%; Pred. No. 31;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 DINPAWYAG 11
 :|||
 DB 273 DIDPAWLAG 281
 :|||

RESULT 11
 F91173
 probable transporter ECs4358 [imported] - Escherichia coli (strain O157:H7, substrain RI
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
 C:Accession: F91173
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: F91173
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-375 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA37781.1; PID:gl3363832; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: ECs4358

Query Match 53.8%; Score 42; DB 2; Length 375;
 Best Local Similarity 50.0%; Pred. No. 35;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 PDINPAWYAG 11
|:::|::|
Db 170 PNLDPWFEGG 179

RESULT 12

F86019
probable transporter yhhJ [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: F86019
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: F86019
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-375 <STO>
A:Cross-references: GB:AE005174; NID:G12518177; PIDN:AAG58618.1; GSPDB:GN00145; UWGP:Z48
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yhhJ

Query Match 53.8%; Score 42; DB 2; Length 375;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 PDINPAWYAG 11
|:::|::|
Db 170 PNLDPWFEGG 179

RESULT 13

S47704
hypothetical 41.1K protein (rhsB-pit intergenic region) - Escherichia coli (strain K-12)
N:Alternate names: yhhJ protein
C:Species: Escherichia coli
C>Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 01-Mar-2002
C:Accession: S47704; H65145
R:Plunkett, G.
submitted to the EMBL Data Library, March 1994
A:Reference number: S47666
A:Accession: S47704
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-375 <PLU>
A:Cross-references: EMBL:U00039; NID:G466582; PIDN:AAB18460.1; PID:G466621
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: H65145
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-375 <BLAT>
A:Cross-references: GB:AE000424; GB:U00095; NID:G2367230; PIDN:ANC76510.1; PID:G1789897;
A:Experimental source: strain K-12, substrain M61655
C:Genetics:
A:Gene: yhhJ

Query Match 53.8%; Score 42; DB 2; Length 375;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 PDINPAWYAG 11
|:::|::|
Db 170 PNLDPWFEGG 179

RESULT 14

T49717
related to BCS1 protein precursor [imported] - Neurospora crassa
N:Alternate names: protein B23L21.300
C:Species: Neurospora crassa
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T49717
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25022
A:Accession: T49717
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-779 <SCH>
A:Cross-references: EMBL:AL356172; GSPDB:GN00116; NCSP:B23L21.300
A:Experimental source: BAC clone B23L21, strain OR74A
C:Genetics:
A:Gene: NCSP:B23L21.300
A:Map position: 6

Query Match 53.2%; Score 41.5; DB 2; Length 779;
Best Local Similarity 56.2%; Pred. No. 86;
Matches 9; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 TPDINPA---WYAGRG 13
|:::|::|
Db 286 TDYINPATRRWYANRG 301

RESULT 15

B72621
hypothetical protein APE1427 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: B72621
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ku
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyru
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: B72621
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-126 <KAW>
A:Cross-references: DDBJ:AP000061; NID:G5104921; PIDN:BAA80424.1; PID:G5105110
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1427
C:Superfamily: Aeropyrum pernix hypothetical protein APE1427

Query Match 52.6%; Score 41; DB 2; Length 126;
Best Local Similarity 60.0%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PDINPAWYAG 11
|:::|::|
Db 22 PDLGEAWYRG 31

Search completed: August 12, 2004, 14:50:25
Job time : 6.74419 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 3.40116 Seconds
(without alignments)
199.024 Million cell updates/sec

Title: US-09-700-643A-1_COPY_12_24
Perfect score: 78
Sequence: 1 TPDINPAWYAGRG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	100.0	98	1 PRRP_BOVIN	P81264 bos taurus
2	74	94.9	83	1 PRRP_RAT	P81277 rattus norv
3	72	92.3	87	1 PRRP_HUMAN	P81277 homo sapien
4	43	55.1	511	1 CP12_CANFA	P56592 canis famil
5	42	53.8	332	1 ISPH_MYCLE	Q9x781 mycobacteri
6	42	53.8	374	1 YHHU_ECOLI	P31993 escherichia
7	41	52.6	313	1 SPBE_STRCL	P37819 streptomyce
8	41	52.6	391	1 GAT5_CHICK	P43692 gallus gall
9	41	52.6	476	1 YAAJ_ECOLI	P11528 escherichia
10	41	52.6	986	1 CYGR_ARBPU	P11528 arabacia pun
11	40	51.3	246	1 NPD_DEIRA	Q9ryd4 deinococcus
12	40	51.3	269	1 XT24_ARATH	P24806 a xylogluca
13	40	51.3	302	1 HEM6_RALSO	Q8xxc3 ralstonia s
14	40	51.3	413	1 EX7L_COREF	Q8fqp1 corynebacte
15	40	51.3	417	1 EX7L_CORGL	Q8xrm3 corynebacte
16	40	51.3	450	1 TD3L_MOUSE	Q8ko07 mus musculu
17	40	51.3	652	1 DREE_CHICK	P1B302 gallus gall
18	40	51.3	678	1 AAV1_RHOSH	Q9jra46 rhodobacter
19	40	51.3	757	1 IMMT_MOUSE	Q8caq8 mus musculu
20	40	51.3	1002	1 ODOI_HUMAN	Q02218 homo sapien
21	39.5	50.6	384	1 GUN_XANAC	P56935 xanthomonas
22	39.5	50.6	482	1 R167_YEAST	P39743 saccharomyc
23	39	50.0	333	1 VINT_BMD2	Q38361 mycobacteri
24	39	50.0	333	1 VINT_BPMER	P25426 mycobacteri
25	39	50.0	371	1 VINT_BPM15	P22884 mycobacteri
26	39	50.0	626	1 THIC_PSEPK	Q88da5 pseudomonas
27	38.5	49.4	240	1 PLSC_HELPY	Q25903 helicobacte
28	38	48.7	197	1 I18B_HUMAN	O95998 homo sapien
29	38	48.7	208	1 YK69_CABEL	P34347 caenorhabdi
30	38	48.7	322	1 GRP2_MOUSE	O89100 m grb2-rela
31	38	48.7	322	1 Y496_METJA	Q57910 methanococc
32	38	48.7	328	1 CNA4_HUMAN	Q9rvr5 homo sapien
33	38	48.7	330	1 GRP2_HUMAN	C75791 h grb2-rela

34	38	48.7	364	1 YAIW_ECOLI	P77562 escherichia
35	38	48.7	376	1 OPS1_LIMPO	P35360 limulus pol
36	38	48.7	376	1 OPS2_LIMPO	P35361 limulus pol
37	38	48.7	455	1 Y955_MYCTU	P71555 mycobacteri
38	38	48.7	512	1 UB93_SCHPO	O94269 schizosacch
39	38	48.7	789	1 CADA_CHICK	P79995 gallus gall
40	38	48.7	875	1 RPP3_RAT	P97675 r ectonucle
41	37	47.4	91	1 RL31_SYNEL	Q8dmk6 synecchococ
42	37	47.4	212	1 PCP_STAAU	Q83596 staphylococ
43	37	47.4	307	1 COBS_CORGL	Q8nnj7 corynebacte
44	37	47.4	321	1 ISPH_VIBPA	Q87887 vibrio para
45	37	47.4	325	1 ISPH_CORGL	Q8nrm2 corynebacte

ALIGNMENTS

RESULT 1

PRRP_BOVIN STANDARD; PRT; 98 AA.
AC P81264;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-releasing peptide PrRP20].
DE hormone)
CN PRH.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.
RC TISSUE=Brain;
RX MEDLINE=98268781; PubMed=9607765;
RA Hiruma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.,
RA "A prolactin-releasing peptide in the brain."
RL Nature 393:272-276(1998).
CC -!- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.
CC -!- TISSUE SPECIFICITY: Medulla oblongata and hypothalamus.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL; AB015417; BAA29025.1;
KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.
FT SIGNAL 1 22
FT PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRRP31.
FT PEPTIDE 33 53 PROLACTIN-RELEASING PEPTIDE PRRP20.
FT PROPEP 58 98
FT MOD RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).
SQ SEQUENCE 98 AA; 10544 MW; 08AC35A13B0FA908 CRC64;

Query Match 100.0%; Score 78; DB 1; Length 98;
Best Local Similarity 100.0%; Pred. No. 3.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRG 13

Db 34 TPDINPAWYAGRG 46

```

RESULT 2
PRRP_RAT
ID_PRRP_RAT STANDARD; PRT; 83 AA.
AC P81278; Q8K3Y0;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing
DE hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-
DE releasing peptide PrRP20].
GN PRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]_
RN SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=98268781; PubMed=9607765;
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
RT "A prolactin-releasing peptide in the brain.";
RL Nature 393:272-276(1998).
RN [2]
RN SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=Sprague-Dawley; TISSUE=Hypothalamus;
RA Anderson S.T., Kokay I.C., Lang T., Grattan D.R., Curlewis J.D.;
RT "Quantitation of prolactin-releasing peptide (PrRP) mRNA expression in
RT specific brain regions during the rat oestrous cycle and in
RT lactation.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RN TISSUE SPECIFICITY.
RX MEDLINE=99426652; PubMed=10498338;
RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,
RA Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,
RA Sumino Y., Fujino M.;
RT "Tissue distribution of prolactin-releasing peptide (PrRP) and its
RT receptor.";
RL Regul. Pept. 83:1-10(1999).
CC -!- FUNCTION: Stimulates prolactin (PRL) release and regulates the
CC expression of prolactin through its receptor GPR10. May stimulate
CC lactotrophs directly to secrete PRL.
CC -!- ALTERNATIVE PRODUCTS.
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P81278-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P81278-2; Sequence=VSP_004370;
CC -!- TISSUE SPECIFICITY: Widely expressed, with highest levels in
CC medulla oblongata and hypothalamus.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AB015418; BAA29026.1; -.
CC DR EMBL; AF521930; AAM82154.1; -.
CC DR PIR; JC7607; JC7607.
CC KW Hormone; Amidation; Signal; Cleavage on pair of basic residues;
CC Alternative splicing.
CC SIGNAL 1 21 BY SIMILARITY.
CC FT PEPTIDE 22 52
CC FT PEPTIDE 33 52 PROLACTIN-RELEASING PEPTIDE PRRP31.
CC FT PROPEP 57 83 PROLACTIN-RELEASING PEPTIDE PRRP20.
CC FT MOD RES 52 52
CC FT VARSPLIC 33 83
CC
CC Query Match 92.3%; Score 72; DB 1; Length 87;
CC Best Local Similarity 92.3%; Pred. No. 2.8e-05;
CC Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

isoform 2).
/FTID=VSP_004370.
/DOC75A264EE4F29 CRC64;
FT
FT
SQ SEQUENCE 83 AA; 9215 MW; 94.9%; Score 74; DB 1; Length 83;
Best Local Similarity 92.3%; Pred. No. 1.2e-05;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 TPDINPAWYAGRG 13
Db 33 TPDINPAWYTGGR 45
RESULT 3
PRRP_HUMAN
ID_PRRP_HUMAN STANDARD; PRT; 87 AA.
AC P81277;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing
DE hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-
DE releasing peptide PrRP20].
GN PRH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RN SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98268781; PubMed=9607765;
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
RT "A prolactin-releasing peptide in the brain.";
RL Nature 393:272-276(1998).
RN [2]
RN TISSUE SPECIFICITY.
RX MEDLINE=99426652; PubMed=10498338;
RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,
RA Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,
RA Sumino Y., Fujino M.;
RT "Tissue distribution of prolactin-releasing peptide (PrRP) and its
RT receptor.";
RL Regul. Pept. 83:1-10(1999).
CC -!- FUNCTION: Stimulates prolactin (PRL) release and regulates the
CC expression of prolactin through its receptor GPR10. May stimulate
CC lactotrophs directly to secrete PRL.
CC -!- TISSUE SPECIFICITY: Medulla oblongata and hypothalamus.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AB015419; BAA29027.1; -.
CC DR EMBL; 602663; -.
CC DR GO; GO:0005180; F:peptide hormone; TAS.
CC KW Hormone; Amidation; Signal.
CC SIGNAL 1 22 BY SIMILARITY.
CC FT PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRRP31.
CC FT PEPTIDE 34 53 PROLACTIN-RELEASING PEPTIDE PRRP20.
CC FT PROPEP 58 87
CC FT MOD RES 53 53
CC FT SEQUENCE 87 AA; 9639 MW; 92.3%; Score 72; DB 1; Length 87;
CC
CC Query Match 92.3%; Score 72; DB 1; Length 87;
CC Best Local Similarity 92.3%; Pred. No. 2.8e-05;
CC Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 TPDINPAWYAGRG 13
|||
DB 34 TPDINPAWYAGRG 46

RESULT 4
CP12 CANFA
ID CP12 CANFA STANDARD; PRT; 511 AA.
AC P56592;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 1A2 (EC 1.14.14.1) (CYPIA2) (DAH2) (Cytochrome P450-
D2).
DE D2).
GN CYP1A2.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE OF 9-511 FROM N.A.
RC STRAIN=Beagle; TISSUE=Liver;
EX MEDLINE=91042464; PubMed=2122230;
RA Uchida T., Komori M., Kitada M., Kamataki T.;
RT Isolation of cDNAs coding for three different forms of liver
RT microsomal cytochrome P-450 from polychlorinated biphenyl-treated
RT beagle dogs."
RL Mol. Pharmacol. 38:644-651(1990).
RN [2]
RP SEQUENCE OF 1-16.
RC STRAIN=Beagle; TISSUE=Liver;
RX MEDLINE=89087526; PubMed=2910310;
RA Ohta K., Motoya M., Komori M., Miura T., Kitada M., Kamataki T.;
RT "A novel form of cytochrome P-450 in beagle dogs. P-450-D3 is a low
RT spin form of cytochrome P-450 but with catalytic and structural
RT properties similar to P-450d."
RL Biochem. Pharmacol. 38:91-96(1989).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- TISSUE SPECIFICITY: Constitutively expressed in liver.
CC -!- INDUCTION: By polychlorinated biphenyl (PCB) in liver and kidney.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR PIR: B3722; B3722.
DR HSP: P00179; 1DT6.
DR InterPro: IPR001128; Cytochrome_P450.
DR InterPro: IPR008066; EP450A.
DR Pfam: PF00667; p450; 1.
DR PRINTS: PR01683; EP450ICYPIA.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT INIT_MET 0
FT METAL 453 453 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 511 AA; 57505 MW; 200904C54F4B3CE7 CRC64;

Query Match 55.1%; Score 43; DB 1; Length 511;
Best Local Similarity 58.3%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGR 12
|||
DB 121 SPDSGFWAAGR 132

RESULT 5

```

```

ISPH_MYCLE
ID ISPH_MYCLE STANDARD; PRT; 332 AA.
AC Q9X781;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (EC 1.17.1.2).
GN ISPH OR LYTB OR LYTB OR MLI938 OR MLCB1222.06C.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus."
RL Nature 409:1007-1011(2001).
CC -!- FUNCTION: Converts 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate
CC into isopentenyl diphosphate (IPP) and dimethylallyl diphosphate
CC (DMAPP) (By similarity).
CC -!- CATALYTIC ACTIVITY: Isopentenyl diphosphate + NAD(P)(+) + H(2)O =
CC (E)-4-hydroxy-3-methylbut-2-en-1-yl diphosphate + NAD(P)H.
CC -!- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; seventh
CC (last) step.
CC -!- SIMILARITY: Belongs to the isph family.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; AL049491; CAB39812.1; ALT_INIT.
CC EMBL; AL583923; CAB30893.1; ALT_INIT.
CC Leproma; MLI938; 1.
CC HAWAP; MF_00191; 1.
CC InterPro: IPR003451; LytB.
CC Pfam; PF02401; LYTB; 1.
CC TIGRFAMs; TIGR00216; isph_lytB; 1.
KW Isoprene biosynthesis; Complete proteome; Oxidoreductase; NADP.
SQ SEQUENCE 332 AA; 36225 MW; F2A82A06EEC35E13 CRC64;

Query Match 53.8%; Score 42; DB 1; Length 332;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 DINPAWYAG 11
|||
DB 270 DIDPAWLAG 278

RESULT 6
YHHJ_ECOLI
ID YHHJ_ECOLI STANDARD; PRT; 374 AA.
AC P31993; P76703;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Hypothetical protein yhhj.
GN YHHJ OR B3485 OR SF3501 OR S4262.
OS Escherichia coli, and

```

OS Shigella flexneri.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562, 623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=K12 / MG1655;
 RX MEDLINE=94316500; PubMed=8041620;
 RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner P.R.;
 RT "Analysis of the Escherichia coli genome. V. DNA sequence of the
 RT region from 76.0 to 81.5 minutes."
 RL Nucleic Acids Res. 22:2576-2586(1994).
 RN [2]
 RP SEQUENCE OF 225-374 FROM N.A.
 RC SPECIES=E.coli; STRAIN=K12;
 RX MEDLINE=93259920; PubMed=8387990;
 RA Zhao S., Sandt C.H., Feulner G., Vlazny D.A., Gray J.A., Hill C.W.;
 RT "Rns elements of Escherichia coli K-12: complex composites of shared
 RT and unique components that have different evolutionary histories."
 RL J. Bacteriol. 175:2799-2808(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
 RX MEDLINE=2272406; PubMed=12384590;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
 RT through comparison with genomes of Escherichia coli K12 and O157."
 RL Nucleic Acids Res. 30:4432-4441(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
 RX MEDLINE=22590274; PubMed=12704152;
 RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
 RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
 RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
 RA Schwartz D.C., Blattner P.R.;
 RT "Complete genome sequence and comparative genomics of Shigella
 RT flexneri serotype 2a strain 2457T."
 RL Infect. Immun. 71:2775-2786(2003).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (Potential).
 CC -!- SIMILARITY: Belongs to the ABC-2 integral membrane protein family.
 CC -!- SIMILARITY: TO E.COLI YBHR AND YBHS AND TO THE C-TERMINAL OF
 CC E.COLI YHIG.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U00039; AB18460.1; ALT_INIT.
 CC EMBL; AE000424; AAC76510.1; ALT_INIT.
 CC EMBL; L02370; AAC61886.1; -.
 CC EMBL; AE015358; AAN44960.1; ALT_INIT.
 CC EMBL; AE016992; AAP19222.1; -.
 CC EcoGene; EGI1767; ynhJ.
 CC InterPro; IPR000412; ABC_transp2.
 CC PROSITE; PS00890; ABC2_MEMBRANE; 1.
 CC Hypothetical protein; Transport; Transmembrane; Inner membrane;
 CC Complete proteome.
 CC
 CC TRANSMEM 23 43 POTENTIAL.
 CC TRANSMEM 173 193 POTENTIAL.
 CC TRANSMEM 230 250 POTENTIAL.
 CC TRANSMEM 256 276 POTENTIAL.
 CC TRANSMEM 284 304 POTENTIAL.
 CC TRANSMEM 343 363 POTENTIAL.

SO SEQUENCE 374 AA; 41061 MW; 02895FB13F493391 CRC64;
 Query Match 53.8%; Score 42; DB 1; Length 374;
 Best Local Similarity 50.0%; Pred. No. 12;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 PDINDAWYAG 11
 Db 169 ENLDPWFEG 178
 [:::|::|:]
 RESULT 7
 ID SPBB_STRCL STANDARD; PRT; 313 AA.
 AC P37819; P72400;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Possible agmatinase (EC 3.5.3.11) (Agmatine ureohydrolase) (AUH)
 DE (Proclavamnic acid amidino hydrolase).
 GN PAH.
 OS Streptomyces clavuligerus.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1901;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 27064 / DSM 738 / NRRL 3585;
 RX MEDLINE=94374706; PubMed=8088547;
 RA Aidoo K.A., Wong A., Alexander D.C., Rittammer R.A.R., Jensen S.E.;
 RT "Cloning, sequencing and disruption of a gene from Streptomyces
 RT clavuligerus involved in clavulanic acid biosynthesis."
 RL Gene 147:41-46(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC Hodgson J.E., Fosberry A., Rawlinson N.S., Ross H.N.M., Neal R.J.,
 RA Arnell J.C., Earl A.J., Lawlor E.J.;
 RA Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: Agmatine + H(2)O = putrescine + urea.
 CC -!- COFACTOR: Manganese (Potential).
 CC -!- PATHWAY: Clavulanic acid biosynthesis.
 CC -!- SIMILARITY: Belongs to the arginase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U87786; AAA62451.1; -.
 CC EMBL; X84101; CAA58904.1; -.
 CC PIR; S57669; S57669
 CC PDB; 1G06; 26-JUN-03.
 CC PDB; 1G07; 26-JUN-03.
 CC InterPro; IPR005925; Agmatinase.
 CC InterPro; IPR006035; Arg agm form.
 CC InterPro; IPR005924; Arginase.
 CC Pfam; PF00491; arginase.1.
 CC PRINTS; PR00116; ARGINASE.
 CC TIGRAME; TIGR01230; agmatinase; 1.
 CC PROSITE; PS00147; ARGINASE_1; 1.
 CC PROSITE; PS00148; ARGINASE_2; 1.
 CC PROSITE; PS01053; ARGINASE_3; 1.
 CC Hydrolase; Manganese; 3D-structure.
 CC METAL 121 121 MANGANESE 1 (BY SIMILARITY).
 CC METAL 144 144 MANGANESE 1 AND 2 (BY SIMILARITY).
 CC METAL 146 146 MANGANESE 2 (BY SIMILARITY).
 CC METAL 148 148 MANGANESE 1 (BY SIMILARITY).
 CC METAL 235 235 MANGANESE 1 AND 2 (BY SIMILARITY).
 CC METAL 237 237 MANGANESE 2 (BY SIMILARITY).
 CC SEQUENCE 313 AA; 33401 MW; 759E9B5644B88D5E CRC64;


```
Query Match 52.6%; Score 41; DB 1; Length 313;
Best Local Similarity 70.0%; Pred. No. 15;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 DINFAYAGR 12
DB 148 DTNFAFYGR 157

RESULT 8
GAT5_CHICK STANDARD; PRT; 391 AA.
AC P43692;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription factor GATA-5 (GATA binding factor-5).
GN GATA5.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn;
RX MEDLINE=94365018; PubMed=8083222;
RA Laverriere A.C., Macneil C., Mueller C., Foellmann R.E.,
RA Burch J.B.E., Evans T.;
RT "GATA-4/5/6, a subfamily of three transcription factors transcribed
RT in developing heart and gut.";
RL J. Biol. Chem. 269:23177-23184(1994).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: More abundant in stomach and small intestine,
CC lower levels in heart, lung and spleen. Very low levels in liver
CC and ovary.
CC -!- DEVELOPMENTAL STAGE: Initially transcribed in the cardiac crescent
CC prior to formation of the primordial heart tube. Following
CC formation of the primitive heart, present in both endocardium and
CC myocardium as well as in other lateral plate derivatives. Also
CC transcribed in the primitive embryonic gut and in late stage
CC embryos is sequentially up-regulated in distinct segments of
CC gastrointestinal epithelia as they undergo terminal
CC differentiation.
CC -!- SIMILARITY: Contains 2 GATA-type zinc fingers.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U1188; AAA57504.1; -
CC FIR; I50702; I50702.
CC HSSP; P17679; IGNF.
CC TRANSFAC; T02683; -
CC InterPro; IPR008013; GATA-N.
CC InterPro; IPR000679; Znf GATA.
CC Pfam; PF05349; GATA-N; 1.
CC Pfam; PF00320; GATA; 2.
CC PRINTS; PRO0619; GATAZNFINGER.
CC SMART; SM00401; ZNF_GATA; 2.
CC PROSITE; PS00344; GATA_ZN_FINGER_1; 2.
CC PROSITE; PS01114; GATA_ZN_FINGER_2; 2.
CC Transcription regulation; Activator; DNA-binding; Zinc-finger;
KW Nuclear protein.
FT ZN_FING 186 210 GATA-TYPE 1.
FT ZN_FING 239 263 GATA-TYPE 2.
SQ SEQUENCE 391 AA; 41858 MW; EF9A283111824260 CRC64;

Query Match 52.6%; Score 41; DB 1; Length 391;
Best Local Similarity 54.5%; Pred. No. 19;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 TPDINAWYAG 11
DB 139 TPFLPSPWTAG 149

RESULT 9
YAAJ_ECOLI STANDARD; PRT; 476 AA.
AC P30143;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative transporter yaaJ.
GN YAAJ OR B0307.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=92334977; PubMed=1630901;
RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
RA Isono K., Mizobuchi K., Nakata A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of
RT the 0-2.4 min region.";
RL Nucleic Acids Res. 20:3305-3308(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1235-1244(1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -!- SIMILARITY: BELONGS TO THE SODIUM:ALANINE SYMPORTER FAMILY
CC (SAF). STRONG, TO H.INFLUENZAE HI0183.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D10483; BAB96585.1; -
CC EMBL; AE000111; AAC73118.1; -
CC FIR; G64720; G64720.
CC EcoGene; EG11555; yaaJ
CC InterPro; IPR002293; AA/rel_permease1.
CC InterPro; IPR001463; Na/Ala_symport.
CC Pfam; PF01235; Na_Ala_sym; 1.
CC PRINTS; PRO0175; NAALASMPORT.
CC TIGRFAMs; TIGR00835; agcs; 1.
CC PROSITE; PS00873; NA_ALANINE_SYMP; 1.
CC Hypothetical protein; Transmembrane; Inner membrane; Transport;
KW Symport; Complete proteome.
FT TRANSMEM 4 24 POTENTIAL.
FT TRANSMEM 81 101 POTENTIAL.
FT TRANSMEM 141 161 POTENTIAL.
FT TRANSMEM 174 194 POTENTIAL.
FT TRANSMEM 207 227 POTENTIAL.
FT TRANSMEM 233 253 POTENTIAL.
FT TRANSMEM 300 320 POTENTIAL.
```


RESULT 12

XT24 ARATH
 ID XT24 ARATH STANDARD; PRT; 259 AA.
 AC P24806; O64956; Q39148; Q39149; Q41904; Q8LDQ0;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Xylglucan endoransglucosylase/hydrolase protein 24 precursor
 DE (EC 2.4.1.207) (At-XTH24) (XTH-24) (Meristem protein 5) (MER1-5
 DE protein) (MERIS protein) (Endo-xylglucan transferase) (Xylglucan
 DE endo-1,4-beta-D-glucanase).
 GN XTH24 OR MER1-5 OR MER15B OR SEN4 OR AT4G30270 OR F9N11.120.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 [1]
 RN SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RP MEDLINE=93005704; PubMed=1840916;
 RX Medford J.I., Elmer J.S., Klee H.J.;
 RA "Molecular cloning and characterization of genes expressed in shoot
 RT apical meristems";
 RL Plant Cell 3:359-370(1991).
 [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=cv. Columbia;
 RC Kanimai T., Tomita E., Nishitani K.;
 RA Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 [3]
 RN SEQUENCE FROM N.A.
 RP STRAIN=cv. Columbia;
 RC MEDLINE=20083488; PubMed=10617198;
 RX Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Dueserhoeft A., Striekema W., Entlan K.-D., Terry N.,
 RA Harris B., Anorge W., Brandt P., Grivell L.A., Rieger M., Mueller M.,
 RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Schmidheini I.,
 RA Kreis M., Delsenij M., Puigdomenech P., Watson M., Schmidheini I.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA Van der Schueren J., Grymoprez B., Chuang Y.-J., Vandenbussche F.,
 RA Braeken M., Weijens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Weitzenecker T., Bothe G., Ransperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
 RA Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
 RA Bernier S., Hempel S., Feldpausch M., Lamberth S., Van den Daale H.,
 RA De Keyser A., Buysschaert C., Gielen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIlroy K., Mayes R.,
 RA Pettett A., Rajandream M.A., Lyne M., Grimm M., Loehner T.-H.,
 RA Borkova D., Bloecker H., Schaefer M., Schaefer A., Mueller-Auer S.,
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Dauner D., Herzl A.,
 RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauver D., Herzi A.,
 RA Neumann S., Argirio A., Vitale D., Liguori R., Piravandi E.,
 RA Massenet O., Quigley F., Clabaud G., Muendlein A., Felber R.,
 RA Schrabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Chetkov F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Farnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threiden J.,
 RA Stoneking T., Kallicki J., Graves T., Harmon G., Edwards J.,
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Minx P., Bentley D., Fulcon B., Miller N., Greco T., Kemp K.,
 RA Kramer J., Fulcon L., Mardis E., Dante M., Pepin K., Hillier L.W.,
 RA Nelson J., Spetch J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,

RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,
 RA Ma P., Zhang J., Preston R., Vill D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Granat S., Shohdy N.A., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
 RA Chen E., Marra M.A., Martienssen R., McCombie W.R.;
 RT "Sequence and analysis of chromosomes 4 of the plant Arabidopsis
 RT thaliana";
 RL Nature 402:769-777(1999).
 [4]
 RN SEQUENCE FROM N.A.
 RP STRAIN=cv. Columbia;
 RC MEDLINE=22954850; PubMed=14593172;
 RX Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J., Cheuk R.F.,
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Yu G.,
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
 RA Arakawa T., Banh J., Banno F., Bowser I., Brooks S.Y., Carninci P.,
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
 RA Hayashizaki Y., Johnson-Hopson C., Huan W., Iida K., Karnes M.,
 RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
 RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
 RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RT "Empirical analysis of transcriptional activity in the Arabidopsis
 RT genome";
 RL Science 302:842-846(2003).
 [5]
 RN SEQUENCE FROM N.A.
 RP Prover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.A.;
 RL "Full-length cDNA from Arabidopsis thaliana";
 RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 [6]
 RN SEQUENCE OF 1-132 FROM N.A.
 RP STRAIN=cv. Columbia; TISSUE=Green siliques;
 RC MEDLINE=94108489; PubMed=8281187;
 RX Hoeft H.R., Desprez T., Amselem J., Chisapello H., Rouze P.,
 RA Caboche M., Molgan A., Jourjon M.-F., Charpentier J.-L.,
 RA Barthomeu P., Guerrier D., Gitaudat J., Quigley F.,
 RA Thomas F., Yu D.-Y., Mache R., Raynal M., Cooke R., Grellet F.,
 RA Delisney M., Parentier Y., de Marcillac G., Gigot C., Fleck J.,
 RA Philipps G., Axelos M., Bardet C., Tremouaygue D., Lescure B.;
 RA "An inventory of 1152 expressed sequence tags obtained by partial
 RT sequencing of cDNAs from Arabidopsis thaliana";
 RL Plant J. 4:1051-1061(1993).
 [7]
 RN SEQUENCE OF 1-120 FROM N.A.
 RP STRAIN=cv. Columbia; TISSUE=Leaf;
 RC MEDLINE=98278374; PubMed=9617812;
 RX Park J.-H., Oh S.A., Kim Y.H., Woo H.R., Nam H.G.;
 RA "Differential expression of senescence-associated mRNAs during leaf
 RT senescence induced by different senescence-inducing factors in
 RL Arabidopsis";
 RL Plant Mol. Biol. 37:445-454(1998).
 [8]
 RN SEQUENCE OF 104-269 FROM N.A.
 RP STRAIN=cv. Landsberg erecta;
 RC MEDLINE=95359399; PubMed=7632911;
 RX Arrowsmith D.A., De Silva J.;
 RA "Characterisation of two tomato fruit-expressed cDNAs encoding
 RT xylglucan endo-transglucosylase";
 RL Plant Mol. Biol. 28:391-403(1995).
 [9]
 RN ENZYMIC ACTIVITY, AND N-GLYCOSYLATION.
 RP MEDLINE=99334624; PubMed=10406121;
 RX Campbell P., Braam J.;
 RA "In vitro activities of four xyloglucan endotransglucosylases from
 RT Arabidopsis";
 RL Plant J. 18:371-382(1999).
 [10]
 RN TISSUE SPECIFICITY.
 RP MEDLINE=21530286; PubMed=11673616;


```

RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens.";
RL Genome Res. 13:1572-1579 (2003).
CC -!- FUNCTION: Bidirectionally degrades single-stranded DNA into large
CC acid-insoluble oligonucleotides, which are then degraded further
CC into small acid-soluble oligonucleotides (By similarity).
CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.
CC -!- SUBUNIT: Heterooligomer composed of large and small subunits (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the xseA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP005217; BAC17888.1; ALT_INIT.
CC HAMAP; MF_00378; -; 1.
CC InterPro; IPR003753; Exonuc_VII_L.
CC InterPro; IPR008994; Nucleic_acid_OB.
CC Pfam; PF02601; Exonuc_VII_L; 1.
CC TIGRFAMs; TIGR00237; xseA; 1.
CC Hydrolase; Nuclease; Exonuclease; Complete proteome.
CC KW SEQUENCE 413 AA; 45171 MW; D3ABD9687C6E5D CRC64;
CC
CC Query Match 51.3%; Score 40; DB 1; Length 413;
CC Best Local Similarity 87.5%; Pred. No. 29;
CC Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 6 PAFYAGRG 13
CC Db 90 PAFYAGRG 97
CC
CC RESULT 15
CC EX7L_CORGL STANDARD; PRT; 417 AA.
CC ID EX7L_CORGL STANDARD; PRT; 417 AA.
CC AC G8NRW3;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
CC DE (Exonuclease VII large subunit).
CC XSEA OR CG11025.
CC OS Corynebacterium glutamicum (Brevibacterium flavum).
CC OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
CC OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
CC OX NCBI_TaxID=1718;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
CC RA Nakagawa S.;
CC RL "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
CC Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Bidirectionally degrades single-stranded DNA into large
CC acid-insoluble oligonucleotides, which are then degraded further
CC into small acid-soluble oligonucleotides (By similarity).
CC -!- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.
CC -!- SUBUNIT: Heterooligomer composed of large and small subunits (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the xseA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP005217; BAC17888.1; ALT_INIT.
CC HAMAP; MF_00378; -; 1.
CC InterPro; IPR003753; Exonuc_VII_L.
CC InterPro; IPR008994; Nucleic_acid_OB.
CC Pfam; PF02601; Exonuc_VII_L; 1.
CC TIGRFAMs; TIGR00237; xseA; 1.
CC Hydrolase; Nuclease; Exonuclease; Complete proteome.
CC KW SEQUENCE 413 AA; 45171 MW; D3ABD9687C6E5D CRC64;
CC
CC Query Match 51.3%; Score 40; DB 1; Length 417;
CC Best Local Similarity 87.5%; Pred. No. 29;
CC Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 6 PAFYAGRG 13
CC Db 94 PAFYAGRG 101
CC
CC Search completed: August 12, 2004, 14:44:51
CC Job time : 4.40116 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 17,9128 Seconds

(without alignments)
228,984 Million cell updates/sec

Title: US-09-700-643A-1_COPY_12_24

Perfect score: 78

Sequence: 1 TPDINPAWYAGRG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTRMBL_25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	98	Q8WN12	Q8WN12 ovis aries
2	55	70.5	117	Q9W624	Q9W624 carassius a
3	53	67.9	54	Q7Z6Y1	Q7Z6Y1 homo sapien
4	53	67.9	465	O60687	O60687 homo sapien
5	53	67.9	465	Q8W85	Q8W85 homo sapien
6	50	64.1	527	O76383	O76383 caenorhabdi
7	48	61.5	457	Q8R054	Q8R054 mus musculu
8	48	61.5	468	Q8K1F8	Q8K1F8 mus musculu
9	47	60.3	430	O27142	O27142 methanobact
10	47	60.3	540	Q886T3	Q886T3 pseudomonas
11	45	57.7	179	Q20170	Q20170 caenorhabdi
12	45	57.7	472	Q91FX2	Q91FX2 chilo iride
13	44	56.4	369	Q98MB6	Q98MB6 rhizobium l
14	44	56.4	379	Q8A1V0	Q8A1V0 bacteroides
15	44	56.4	398	Q9SJR5	Q9SJR5 arabidopsis
16	44	56.4	398	Q8GYS6	Q8GYS6 arabidopsis

17	44	56.4	816	10	Q7XMH8	Q7Xmh8 oryza sativ
18	43	55.1	86	16	Q7U4H3	Q7U4h3 synechococc
19	43	55.1	196	10	Q84WL0	Q84wl0 arabidopsis
20	43	55.1	380	10	Q8L8A7	Q8l8a7 arabidopsis
21	43	55.1	387	16	Q7VUL5	Q7vul5 bordetella
22	43	55.1	419	16	Q8UKP2	Q8ukp2 agrobacteri
23	43	55.1	432	16	Q7WQX8	Q7wqx8 bordetella
24	43	55.1	432	16	Q7W200	Q7w200 bordetella
25	43	55.1	501	13	Q8J136	Q8j136 brachydanio
26	43	55.1	548	10	Q9LFA0	Q9lfa0 arabidopsis
27	43	55.1	605	16	Q8P3N1	Q8p3n1 xanthomonas
28	43	55.1	614	16	Q8PFF24	Q8pf24 xanthomonas
29	43	55.1	630	4	Q9NKS8	Q9nxs8 homo sapien
30	43	55.1	813	17	Q8ZXH2	Q8zxh2 pyrobaculum
31	43	55.1	1595	4	Q8IVG6	Q8ivg6 homo sapien
32	42	53.8	72	8	Q9TD69	Q9td69 mytilus cal
33	42	53.8	120	8	Q85R31	Q85r31 mytilus tro
34	42	53.8	130	16	Q82RV0	Q82rv0 streptomyce
35	42	53.8	211	16	Q8XVH7	Q8xvh7 raistonias
36	42	53.8	341	16	Q7VM68	Q7vm68 haemophilus
37	42	53.8	374	16	Q8ZLE1	Q8zle1 salmonella
38	42	53.8	375	16	Q8X5S4	Q8x5s4 escherichia
39	42	53.8	402	16	Q8FCL8	Q8fcl8 escherichia
40	42	53.8	501	16	Q88S54	Q88s54 pseudomonas
41	42	53.8	505	16	Q88RQ2	Q88rq2 pseudomonas
42	42	53.8	548	2	Q9RMJ1	Q9rmj1 neisseria m
43	42	53.8	548	2	Q9RMJ0	Q9rmj0 neisseria m
44	42	53.8	694	16	Q98DA7	Q98da7 rhizobium l
45	41.5	53.2	779	3	Q9P5J9	Q9p5j9 neurospora

ALIGNMENTS

RESULT 1

Q8WN12 PRELIMINARY; PRT; 98 AA.
 ID Q8WN12
 AC Q8WN12, 2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DB 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Preprolactin-releasing peptide.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Curlewis J.D., Kusters D.H.L., Barclay J.L., Anderson S.T.;
 RT "Prolactin-releasing peptide (PrRP) in the ewe: cDNA cloning, mRNA
 RT distribution and effects on prolactin secretion in vitro and in
 RT vivo.";
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF450453; AAL47178.1;
 SQ SEQUENCE 98 AA; 10513 MW; 2A53331ED62CAAB5 CRC64;

Query March 100.0%; Score 78; DB 6; Length 98;
 Best Local Similarity 100.0%; Pred. No. 5.5e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRG 13
 |||||
 Db 34 TPDINPAWYAGRG 46

RESULT 2

Q9W624 PRELIMINARY; PRT; 117 AA.
 ID Q9W624
 AC Q9W624;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

```

DE C-RF amide.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Satake H., Minakata H., Fujimoto M.;
RT "Carassius RAmide (C-RF amide).";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB020024; BAA76682.1; -.
SQ SEQUENCE 117 AA; 12879 MW; DSDC4CB22038C2B0 CRC64;

Query Match 70.5%; Score 55; DB 13; Length 117;
Best Local Similarity 61.5%; Pred. No. 0.33;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TPDIINPAWYAGRG 13
Db 56 SPEIDPFWYVGRG 68

RESULT 3
Q7Z6Y1 PRELIMINARY; PRT; 54 AA.
ID Q7Z6Y1 AC
AC Q7Z6Y1,
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DJ479J7.3 (Sushi-repeat protein (SRPUL)) (Fragment).
DE DJ479J7.3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lawlor S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL035608; CAB55892.1; -.
FT NON_TER 54
SQ SEQUENCE 54 AA; 6110 MW; E2F3C99F7B961A9F CRC64;

Query Match 67.9%; Score 53; DB 4; Length 54;
Best Local Similarity 61.5%; Pred. No. 0.31;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TPDIINPAWYAGRG 13
Db 18 TPAVPTWYAGSG 30

RESULT 4
O60687 PRELIMINARY; PRT; 465 AA.
ID O60687 AC
AC O60687;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sushi-repeat protein (Sushi-repeat containing protein).
DE SRPUL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kurosawa H., Inukai T., Inaba T., Goi K., Chang K.-S., Sinjo T.,
RA Rakestraw K.M., Naeve C.W., Look T.A.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]

```

RP SEQUENCE FROM N.A.
RA "C.-H., Chen H., Peng J., Chen Y.;
RT Huang and characterization of the sushi-repeat containing protein
RT (SRP) as a novel interaction partner of Rn type C glycoprotein
RL (RNCG).";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF060567; AAC15765.1; -
DR EMBL; AF060567; AAC15765.1; -
DR EMBL; AF393649; AAM73693.1; -
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR003410; Hyalin.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF02494; HYR; 1.
DR Pfam; PF00084; sushi; 3.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
SQ SEQUENCE 465 AA; 52971 MW; 4D752B187FF3EBB8 CRC64;

Query Match 67.9%; Score 53; DB 4; Length 465;
Best Local Similarity 61.5%; Pred. No. 2.9;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRG 13
||| : |||||
Db 18 TPVPTWYAGSG 30

RESULT 5
Q8W85 PRELIMINARY; PRT; 465 AA.
ID AC Q8W85
AC Q8W85; 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Sushi-repeat protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Placenta;
RC Straussberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC020733; AAH20733.1; -
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR003410; Hyalin.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF02494; HYR; 1.
DR Pfam; PF00084; sushi; 3.
DR SMART; SM00032; CCP; 3.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
SQ SEQUENCE 465 AA; 52957 MW; 3D7229487DAB8BD CRC64;

Query Match 67.9%; Score 53; DB 4; Length 465;
Best Local Similarity 61.5%; Pred. No. 2.9;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRG 13
||| : |||||
Db 18 TPVPTWYAGSG 30

RESULT 6
O76383 PRELIMINARY; PRT; 527 AA.
ID AC O76383
AC O76383; 01-NOV-1998 (T-EMBLrel. 08, Created)
DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN C24G6.6.
OS Caenorhabditis elegans.

RT "Complete genome sequence of Methanobacterium thermoautotrophicum
deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
DR EMBL; AE000877; AAB85559.1; 5'.
DR FJ; B69009; B69009.
DR InterPro; IPR002510; Peptidase_U62.
DR Pfam; PF01523; PmbA_Tldd; 1.
KW Complete proteome.
SQ SEQUENCE 430 AA; 46062 MW; C4FAE47C111749E6 CRC64;

Query Match 60.3%; Score 47; DB 17; Length 430;
Best Local Similarity 80.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 DPNPWAYGR 12
Db 191 DPNPWAYGR 200

RESULT 10

Q886T3 PRELIMINARY; PRT; 540 AA.
AC Q886T3 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Methyl-accepting chemotaxis protein.
GN PSPT01493.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000.
RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
RA Berry K., Utterback T., Van Aken S., Feldblyum T., Gwinn M.,
RA Dodson R., Deboy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidsen T.,
RA White O., Fraser C., Collier A.;
RT "Complete sequence of Pseudomonas syringae";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AS016861; AAC5014.1; -.
DR TIGR; PSPT01493; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0006935; P:chemotaxis; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR004089; Chmtaxis_transd.
DR InterPro; IPR003660; HAMF.
DR InterPro; IPR004090; Me_chemotaxis.
DR Pfam; PF00672; HAMF; 1.
DR Pfam; PF00015; MCPsignal; 1.
DR PRINTS; PR00260; CHEMTRNSDUCR.
DR PROSITE; PS0111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS00885; HAMF; 1.
KW Complete proteome.
SQ SEQUENCE 540 AA; 58872 MW; 2B203D4F1B1B8094 CRC64;

Query Match 70.3%; Score 47; DB 16; Length 540;
Best Local Similarity 80.0%; Pred. No. 31;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 DPNPWAYGR 12
Db 150 DPNPWAYGR 159

RESULT 11

Q20170 PRELIMINARY; PRT; 179 AA.
ID Q20170
AC Q20170
DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE C. elegans WRT-3 protein (Corresponding sequence F38E11.7).
DR WRT-3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RL investigating biology.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Matthews P.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z68342; CAA92775.2; -.
SQ SEQUENCE 179 AA; 20965 MW; F9FA6891A2BF2BDD CRC64;

Query Match 57.7%; Score 45; DB 5; Length 179;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PDINPAMY 9
Db 159 PDVKEPAMY 166

RESULT 12

Q91FX2 PRELIMINARY; PRT; 472 AA.
AC Q91FX2;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE 199R.
OS Chilo iridescent virus (CIV) (Insect iridescent virus type 6).
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OX NCBI_TaxID=10488;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H., Darai G., Fluegel R.M.;
RT "DNA analysis of insect iridescent virus 6: evidence for circular
RT permutation and terminal redundancy.";
RL J. Virol. 49:609-614(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86174607; PubMed=3959991;
RA Lorbacher de Ruiz H., Gelderblom H., Hofmann W., Darai G.;
RT "Insect iridescent virus type 6 induced toxic degenerative hepatitis
RT in mice.";
RL Med. Microbiol. Immunol. 175:43-53(1986).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=87321126; PubMed=2820141;
RA Schnitzler P., Soltan J.B., Fischer M., Reissner H., Scholz J.,
RA Delius H., Darai G.;
RT "Molecular cloning and physical mapping of the genome of insect
RT iridescent virus type 6: further evidence for circular permutation of
RT the viral genome.";
RL Virology 160:66-74(1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89073752; PubMed=3201750;
RA Fischer M., Schnitzler P., Delius H., Darai G.;
RT "Identification and characterization of the repetitive DNA element in
RT the genome of insect iridescent virus type 6.";
RL Virology 167:485-496(1988).

RN SEQUENCE FROM N.A.
 RX MEDLINE=9219696; PubMed=1549908;
 RA Handermann M., Schnitzler P., Rosen-Wolff A., Raab K., Sonntag K.C.,
 Darai G.;
 RT "Identification and mapping of origins of DNA replication within the
 RT DNA sequences of the genome of insect iridescent virus type 6.";
 RL Virus Genes 6:19-32(1992).
 RN
 RN SEQUENCE FROM N.A.
 RP MEDLINE=93118242; PubMed=1475907;
 RA Sonntag K.C., Darai G.;
 RT "Characterization of the third origin of DNA replication of the genome
 RT of insect iridescent virus type 6.";
 RL Virus Genes 6:333-342(1992).
 RN
 RN SEQUENCE FROM N.A.
 RP MEDLINE=93260401; PubMed=8492091;
 RA Stohwasser R., Raab K., Schnitzler P., Janssen W., Darai G.;
 RT "Identification of the gene encoding the major capsid protein of
 RT insect iridescent virus type 6 by polymerase chain reaction.";
 RL J. Gen. Virol. 74:873-879(1993).
 RN
 RN SEQUENCE FROM N.A.
 RP MEDLINE=94167241; PubMed=8121799;
 RA Schnitzler P., Hug M., Handermann M., Janssen W., Koonin E.V.,
 Delius H., Darai G.;
 RT "Identification of genes encoding zinc finger proteins, non-histone
 RT chromosomal HMG protein homologue, and a putative GTP phosphohydrolase
 RT in the genome of Chilo iridescent virus.";
 RL Nucleic Acids Res. 22:158-166(1994).
 RN
 RN SEQUENCE FROM N.A.
 RP MEDLINE=94353641; PubMed=8073636;
 RA Sonntag K.C., Schnitzler P., Koonin E.V., Darai G.;
 RT "Chilo iridescent virus encodes a putative helicase belonging to a
 RT distinct family within the 'DEAD/H' superfamily: implications for the
 RT evolution of large DNA viruses.";
 RL Virus Genes 8:151-158(1994).
 RN
 RN SEQUENCE FROM N.A.
 RP MEDLINE=95213160; PubMed=7698884;
 RA Sonntag K.C., Schnitzler P., Janssen W., Darai G.;
 RT "Identification of the primary structure and the coding capacity of
 RT the genome of insect iridescent virus type 6 between the genome
 RT coordinates 0.310 and 0.347 (7990 bp).";
 RL Intervirology 37:287-297(1994).
 RN
 RN SEQUENCE FROM N.A.
 RP MEDLINE=9422906; PubMed=8021587;
 RA Schnitzler P., Sonntag K.C., Muller M., Janssen W., Bugert J.J.,
 Koonin E.V., Darai G.;
 RT "Insect iridescent virus type 6 encodes a polypeptide related to the
 RT largest subunit of eukaryotic RNA polymerase II.";
 RL J. Gen. Virol. 75:1557-1567(1994).
 RN
 RN SEQUENCE FROM N.A.
 RP MEDLINE=98141693; PubMed=9482589;
 RA Bahr U., Tidona C.A., Darai G.;
 RT "The DNA sequence of Chilo iridescent virus between the genome
 RT coordinates 0.101 and 0.391; similarities in coding strategy between
 RT insect and vertebrate iridoviruses.";
 RL Virus Genes 15:235-245(1997).
 RN
 RN SEQUENCE FROM N.A.
 RP MEDLINE=99125223; PubMed=9926400;
 RA Muller K., Tidona C.A., Bahr U., Darai G.;
 RT "Identification of a thymidylate synthase gene within the genome of
 RT Chilo iridescent virus.";
 RL Virus Genes 17:243-258(1998).
 RN
 RN SEQUENCE FROM N.A.
 RP MEDLINE=99383793; PubMed=10456793;

RA Muller K., Tidona C.A., Darai G.;
 RT "Identification of a gene cluster within the genome of Chilo
 RT iridescent virus encoding enzymes involved in viral DNA replication
 RT and processing.";
 RL Virus Genes 18:243-264(1999).
 RN
 RN SEQUENCE FROM N.A.
 RP MEDLINE=21342589; PubMed=11448171;
 RA Jakob N.J., Muller K., Bahr U., Darai G.;
 RT "Analysis of the First Complete DNA Sequence of an Invertebrate
 RT Iridovirus: Coding Strategy of the Genome of Chilo Iridescent Virus.";
 RL Virology 286:182-196(2001).
 RN
 RN SEQUENCE FROM N.A.
 RP Jakob N.J., Mueller K., Bahr U., Darai G.;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF303741; AAK82060.1; -;
 SQ SEQUENCE 472 AA; 52183 MW; 8BDF96D987B8F94B CRC64;
 Query Match 57.7%; Score 45; DB 12; Length 472;
 Best Local Similarity 50.0%; Pred. No. 57;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 TPDINPAWYAGR 12
 Db 298 SPSVNFVWVAGQ 309
 RESULT 13
 Q98MB6 PRELIMINARY; PRT; 369 AA.
 AC Q98MB6;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Probable oxidoreductase.
 GN MLI0645.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Phyllobacteriaceae; Mesorhizobium.
 ON NCBI_TaxID=381;
 RX STRAIN=MAFF303099;
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21082930; PubMed=11214968;
 RA Kato T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 Watanabe A., Igesawa K., Ishikawa A., Kawashima K., Kimura T.,
 Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 DR EMBL: AP002995; BAB48197.1; -;
 DR InterPro: IPR001395; Aldo/ket_red.
 DR Pfam: PF00248; aldo ket red; 1.
 DR PRINTS: PR00069; ALDKETRD7ASE.
 DR ProDom: PD000288; Aldo/ket_red; 2.
 KW Complete Proteome.
 SQ SEQUENCE 369 AA; 39937 MW; 1FA623F2AAC5E907 CRC64;
 Query Match 56.4%; Score 44; DB 16; Length 369;
 Best Local Similarity 100.0%; Pred. No. 64;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 6 PAWYAGR 12
 Db 161 PAWYAGR 167
 RESULT 14
 Q8A1V0 PRELIMINARY; PRT; 379 AA.
 ID Q8A1V0

```

AC Q8A1V0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative endonuclease.
GN BT3558.
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=12663928;
RA Xu J., Bjursell M.K., Hinrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
DR EMBL; AEO16941; AAC078664.1; -.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001604; F:endonuclease.
DR InterPro; IPR000437; Prok_lipoprot_S.
DR Pfam; PF01223; Endonuclease; 1.
DR SMART; SW00477; NUC; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Endonuclease; Complete proteome.
SQ SEQUENCE 379 AA; 41860 MW; 814F8E4D7D3E0FBB CRC64;

Query Match 56.4%; Score 44; DB 16; Length 379;
Best Local Similarity 72.7%; Pred. No. 65;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PDINPANYAGR 12
Db 126 PDPNPSGYAGR 136
|||:|||||
|||:|||||

RESULT 15
Q9SJR5 PRELIMINARY; PRT; 398 AA.
AC Q9SJR5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Atg36400 protein (Transcription activator).
GN AT2G36400 OR GRL3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana";
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.

```

```

RA Kim J.H., Kende H.;
RT "A novel class of transcription activators interacting with putative
co-activators in Arabidopsis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006919; AAC24624.1; -.
DR EMBL; AY102636; AAM52878.1; -.
DR PIR; C84780; C84780.
SQ SEQUENCE 398 AA; 43707 MW; 2A6E15497305AF74 CRC64;

Query Match 56.4%; Score 44; DB 10; Length 398;
Best Local Similarity 63.6%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 PDINPANYAGR 12
Db 130 PHYOPAWYLR 140
|||:|||||
|||:|||||

Search completed: August 12, 2004, 14:49:00
Job time : 19.9128 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 27.5872 Seconds
(without alignments)
133.146 Million cell updates/sec

Title: US-09-700-643A-1_COPY_12_24

Perfect score: 78

Sequence: 1 TPDINPAWYAGRG 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq 29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	19	2 AAW31370	AAW31370 Bovine G
2	78	100.0	19	2 AAW95185	AAW95185 Bovine pi
3	78	100.0	19	6 ABU60830	ABU60830 Peptide p
4	78	100.0	20	2 AAW31374	AAW31374 Bovine G
5	78	100.0	20	2 AAW95191	AAW95191 Bovine pi
6	78	100.0	20	2 AAW97232	AAW97232 Bovine pi
7	78	100.0	20	3 AAY49301	AAY49301 19P2 liga
8	78	100.0	20	3 AAB10350	AAB10350 Bovine ox
9	78	100.0	20	4 AAB90996	AAB90996 Prolactin
10	78	100.0	20	4 AAB46954	AAB46954 Peptide P
11	78	100.0	20	4 AAG62519	AAG62519 Bovine CR
12	78	100.0	20	5 AAR26402	AAR26402 Bovine pr
13	78	100.0	20	6 ABU60834	ABU60834 Peptide p
14	78	100.0	21	2 AAW31375	AAW31375 Bovine G
15	78	100.0	21	2 AAW95192	AAW95192 Bovine pi
16	78	100.0	21	3 AAB10351	AAB10351 Bovine ox
17	78	100.0	21	4 AAG62520	AAG62520 Bovine CR
18	78	100.0	21	6 ABU60835	ABU60835 Peptide p
19	78	100.0	22	2 AAW31376	AAW31376 Bovine G
20	78	100.0	22	2 AAW95193	AAW95193 Bovine pi
21	78	100.0	22	3 AAB10352	AAB10352 Bovine ox
22	78	100.0	22	4 AAG62521	AAG62521 Bovine CR
23	78	100.0	22	6 ABU60836	ABU60836 Peptide p
24	78	100.0	29	2 AAW31369	AAW31369 Bovine G
25	78	100.0	29	2 AAW95184	AAW95184 Bovine pi

26	78	100.0	29	6 ABU60829	ABU60829 peptide p
27	78	100.0	31	2 AAW31371	AAW31371 Bovine G
28	78	100.0	31	2 AAW95188	AAW95188 Bovine pi
29	78	100.0	31	2 AAW87613	AAW87613 Bovine 19
30	78	100.0	31	2 AAW97218	AAW97218 Bovine pi
31	78	100.0	31	3 AAY49298	AAY49298 19P2 liga
32	78	100.0	31	3 AAY49290	AAY49290 19P2 liga
33	78	100.0	31	3 AAB10347	AAB10347 Bovine ox
34	78	100.0	31	4 AAG62516	AAG62516 Bovine CR
35	78	100.0	31	5 AAE26399	AAE26399 Bovine pr
36	78	100.0	31	6 ABU60835	ABU60835 Peptide p
37	78	100.0	31	6 AAW31372	AAW31372 Bovine G
38	78	100.0	32	2 AAW95189	AAW95189 Bovine pi
39	78	100.0	32	2 AAB10348	AAB10348 Bovine ox
40	78	100.0	32	3 AAG62517	AAG62517 Bovine CR
41	78	100.0	32	4 AAG62517	ABU60832 Peptide p
42	78	100.0	32	6 ABU60832	AAW31373 Bovine G
43	78	100.0	33	2 AAW31373	AAW31373 Bovine pi
44	78	100.0	33	2 AAW95190	AAW95190 Bovine G
45	78	100.0	33	3 AAY49297	AAY49297 19P2 liga

ALIGNMENTS

RESULT 1
AAW31370
ID AAW31370 standard; peptide; 19 AA.
XX
AC AAW31370;
XX
DT 06-APR-1998 (first entry)
XX
DE Bovine G protein-coupled receptor ligand fragment from P-2 fraction.
XX
KW G protein-coupled receptor; ligand binding; pharmaceutical; modulator;
KW pituitary; central nervous system; pancreas; prophylactic;
KW therapeutic agent.
XX
OS Bos taurus.
XX
PN WO9724436-A2.
XX
PD 10-JUL-1997.
XX
PF 26-DEC-1996; 96WO-JFP003821.
XX
PR 28-DEC-1995; 95JP-00343371.
PR 15-MAR-1996; 96JP-00059419.
PR 12-AUG-1996; 96JP-00211805.
PR 18-SEP-1996; 96JP-00246573.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;
PI Kitada C;
XX
DR WPI; 1997-363672/33.
XX
N-PSDB; AAV02393.
PT Ligand peptide for G protein-coupled receptor - acts by modulating
function in the central nervous system, pancreas and pituitary gland.
XX
PS Claim 2; Page 160; 258pp; English.

This sequence represents a peptide fragment of a novel bovine pituitary derived ligand corresponding to amino acid residues 34 to 52 of the sequence in AAW31368. This fragment was obtained by purification and analysis of the N-terminal sequence of a P-2 fraction which is used in an assay to monitor ligand binding to the G protein-coupled receptor protein. Pharmaceutical compositions containing this ligand may be used as a pituitary function modulator, a central nervous system modulator or a pancreatic function modulator. This ligand could have specific

CC applications as a prophylactic or therapeutic agent for dementia,
 CC depression, hyperkinetic syndrome, disturbance of consciousness, anxiety
 CC syndrome, schizophrenia, trauma, growth hormone secretory disease, hyper-
 CC and polyphagia, hypercholesterolaemia, hyperglyceridaemia,
 CC hyperlipidaemia, hyperprolactinaemia, diabetes, cancer, pancreatitis,
 CC renal disease, Turner's syndrome, neurosis, rheumatoid arthritis, spinal
 CC injury, transient brain ischaemia, amyotrophic lateral sclerosis, acute
 CC myocardial infarction, spinocerebellar degeneration, bone fracture,
 CC trauma, atopic dermatitis, osteoporosis, asthma, epilepsy, infertility
 CC and/or oligogalactia. Assays can also be developed to screen compounds
 CC which are capable of altering the binding activity of the ligand thus
 CC affecting activation of the G protein-coupled receptor protein
 XX
 SQ Sequence 19 AA;

Query Match 100.0%; Score 78; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 2.6e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRG 13
 |||||
 Db 1 TPDINPAWYAGRG 13

RESULT 2
 AAW95185
 ID AAW95185 standard; peptide; 19 AA.
 AC AAW95185;
 XX
 DT 10-MAR-1999 (first entry)
 XX
 DE Bovine pituitary-derived ligand polypeptide partial sequence.

XX Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
 KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
 KW Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
 KW gene therapy; transgenic animal; bovine.
 XX
 OS Bos sp.
 XX
 PN WO9849295-A1.
 XX
 PD 05-NOV-1998.
 XX
 PF 27-APR-1998; 98WO-JP001923.
 XX
 PR 28-APR-1997; 97JP-00109974.
 XX
 PR (TAKE) TAKEDA CHEM IND LTD.
 XX
 PA Hinuma S, Fukusumi S;
 PI
 XX WPI; 1999-009423/01.
 DR
 XX New polypeptide ligand for orphan G protein coupled receptors - used for
 PT treating disorders of central nervous system, pituitary and pancreas, and
 PT for drug screening.

PS Example 16; Page 145; 206pp; English.
 XX
 CC The invention relates to a murine pituitary-derived ligand polypeptide
 CC which is a ligand for the G-protein coupled orphan receptor designated
 CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing
 CC the ligand polypeptide encoding DNA are used to produce a recombinant
 CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate
 CC function of the pituitary, central nervous system, pancreas and other
 CC tissues and can be used to screen for agents that modulate binding of the
 CC polypeptide to the receptor; to quantify the amount of receptor in a
 CC sample and to raise antibodies. They may also be used therapeutically,

CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's
 CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;
 CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;
 CC rheumatoid arthritis, epilepsy and many others; also to improve post-
 CC operative nutritional status and as vasopressor. Transgenic animals
 CC carrying the ligand polypeptide encoding DNA or its mutant are used to
 CC study the function of the polypeptide-expressing genes, as models of
 CC disease, for drug screening and as source of cell lines. The ligand
 CC polypeptide DNA is used as a source of probes and primers; to identify
 CC related sequences; in receptor-binding assays; for production of Ab and
 CC antisera; in drug development; for gene therapy and to develop transgenic
 CC animals. The present sequence represents a bovine pituitary-derived
 CC ligand polypeptide obtained by N-terminal sequence analysis
 XX
 SQ Sequence 19 AA;

Query Match 100.0%; Score 78; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 2.6e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRG 13
 |||||
 Db 1 TPDINPAWYAGRG 13

RESULT 3
 ABU60830
 ID ABU60830 standard; peptide; 19 AA.
 AC ABU60830;
 XX
 DT 06-MAY-2003 (first entry)
 XX
 DE Peptide production by gene recombination associated peptide #14.

XX Peptide production; low-molecular peptide; KiSS-1; GPR8 ligand;
 KW gene recombination.
 XX
 OS Bos sp.
 XX
 PN WO200292829-A1.
 XX
 PD 21-NOV-2002.
 XX
 PF 16-MAY-2002; 2002WO-JP004735.
 XX
 PR 17-MAY-2001; 2001JP-00147341.
 XX
 PR (TAKE) TAKEDA CHEM IND LTD.
 XX
 PA Nishimura O, Suenaga M, Ito T, Kitada C;
 PI
 XX WPI; 2003-129302/12.
 DR
 XX Process for producing peptides e.g. KiSS-1 peptide and GPR8 ligand for
 PT subsequent applications by gene recombination technique through tandem
 PT repeats to provide precursor protein with specific cleavage sites.

PS Disclosure; Page 60; 87pp; Japanese.
 XX
 CC The invention describes a method of producing a peptide comprising the
 CC excision of the N and C-terminals of a target peptide with enzymes or
 CC chemically through the attached cleavage sites repeated by ligation in a
 CC precursor protein. The method is for producing (low-molecular) peptides
 CC e.g. KiSS-1 peptide and GPR8 ligand for subsequent applications by the
 CC gene recombination technique through tandem repeats to provide a
 CC precursor protein with specific cleavage sites. With this method, peptide
 CC production can be carried out easily to provide large quantities of the
 CC required peptides. This is the amino acid sequence of a peptide
 CC associated with the peptide production method of the invention
 XX
 SQ Sequence 19 AA;

Query Match 100.0%; Score 78; DB 6; Length 19;
 Best Local Similarity 100.0%; Pred. No. 2.6e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRG 13
 |||||
 DB 1 TPDINPAWYAGRG 13

RESULT 4
 AAW31374
 ID AAW31374 standard; peptide; 20 AA.
 AC AAW31374;
 DT 06-APR-1998 (first entry)
 DE Bovine G protein-coupled receptor ligand peptide fragment 4.
 KW G protein-coupled receptor; ligand binding; pharmaceutical; modulator;
 KW pituitary; central nervous system; pancreas; prophylactic;
 KW therapeutic agent.
 KW Bos taurus.
 OS Bos taurus.
 PN WO9724436-A2.
 PD 10-JUL-1997.
 PF 26-DEC-1996; 96WO-JP003821.
 PR 28-DEC-1995; 95JP-00343371.
 PR 15-MAR-1996; 96JP-00059419.
 PR 12-AUG-1996; 96JP-00211805.
 PR 18-SEP-1996; 96JP-00246573.
 PA (TAKE) TAKEDA CHEM IND LTD.
 PI Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;
 Kitada C;
 WPI; 1997-363672/33.
 N-PSDB; AAV02397.

Ligand peptide for G protein-coupled receptor - acts by modulating
 function in the central nervous system, pancreas and pituitary gland.

Claim 2; Page 161; 258pp; English.

This sequence represents a peptide fragment of a novel bovine pituitary
 derived ligand corresponding to amino acid residues 34 to 53 of the
 sequence in AAW31374 and is used in an assay to monitor ligand binding to
 the G protein-coupled receptor protein. Pharmaceutical compositions
 containing this ligand may be used as a pituitary function modulator, a
 central nervous system modulator or a pancreatic function modulator. This
 ligand could have specific applications as a prophylactic or therapeutic
 agent for dementia, depression, hyperkinetic syndrome, disturbance of
 consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone
 secretory disease, hyper and polyphagia, hypercholesterolaemia,
 hyperglycaemia, hyperlipidaemia, hyperprolactinaemia, diabetes,
 cancer, pancreatitis, renal disease, Turner's syndrome, neurosis,
 rheumatoid arthritis, spinal injury, transient brain ischaemia,
 amyotrophic lateral sclerosis, acute myocardial infarction,
 spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 osteoporosis, asthma, epilepsy, infertility and/or oligogalactia. Assays
 can also be developed to screen compounds which are capable of altering
 the binding activity of the ligand thus affecting activation of the G
 protein-coupled receptor protein

Sequence 20 AA;

Query Match 100.0%; Score 78; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.8e-06;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRG 13
 |||||
 DB 1 TPDINPAWYAGRG 13

RESULT 5
 AAW95191
 ID AAW95191 standard; peptide; 20 AA.
 AC AAW95191;
 DT 10-MAR-1999 (first entry)
 DE Bovine pituitary-derived ligand polypeptide fragment.
 KW Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
 KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
 KW Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
 KW gene therapy; transgenic animal; bovine.
 OS Bos sp.
 PN WO9849295-A1.
 PD 05-NOV-1998.
 PF 27-APR-1998; 98WO-JP001923.
 PR 28-APR-1997; 97JP-00109974.
 PA (TAKE) TAKEDA CHEM IND LTD.
 PI Hinuma S, Fukusumi S;
 WPI; 1999-009423/01.
 New polypeptide ligand for orphan G protein coupled receptors - used for
 treating disorders of central nervous system, pituitary and pancreas, and
 for drug screening.

Example 19; Page 151; 206pp; English.

The invention relates to a murine pituitary-derived ligand polypeptide
 which is a ligand for the G-protein coupled orphan receptor designated
 GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing
 the ligand polypeptide encoding DNA are used to produce a recombinant
 ligand polypeptide. The ligand polypeptide, and its fragments, modulate
 function of the pituitary, central nervous system, pancreas and other
 tissues and can be used to screen for agents that modulate binding of the
 polypeptide to the receptor; to quantify the amount of receptor in a
 sample and to raise antibodies. They may also be used therapeutically,
 e.g. to treat senile dementia, Alzheimer's, Parkinson's or Huntington's
 diseases; Creutzfeld-Jakob disease; disorders of growth hormone secretion; cancer;
 diabetes; schizophrenia; epilepsy and many others, also to improve post-
 rheumatoid arthritis, poisoning by heavy metals or drugs;
 operative nutritional status and as vasopressor. Transgenic animals
 carrying the ligand polypeptide encoding DNA or its mutin are used to
 study the function of the polypeptide-expressing genes, as models of
 disease, for drug screening and as source of cell lines. The ligand
 polypeptide DNA is used as a source of probes and primers; to identify
 related sequences; in receptor-binding assays; for production of Ab and
 antisera; in drug development; for gene therapy and to develop transgenic
 animals. The present sequence represents a bovine genome-derived ligand
 polypeptide fragment which is similar to the murine ligand-polypeptide

Sequence 20 AA;

Query Match 100.0%; Score 78; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRG 13
| | | | | | | | | | | | | | |
Db 1 TPDINPAWYAGRG 13

RESULT 6
AAW97232
ID AAW97232 standard; peptide; 20 AA.
XX
AC AAW97232;
XX
DT 06-MAY-1999 (first entry)
XX
DE Bovine pituitary-derived ligand polypeptide fragment.
XX
KW Bovine pituitary-derived ligand; modulation; prolactin secretion;
KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;
KW menopausal syndrome; euthyroid; hypometabolism; lactation;
KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;
KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;
KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;
KW contraceptive; placental function; choriocarcinoma; hydatid mole;
KW interruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;
KW abnormal lipidmetabolism; oxytocia.
XX
OS Bos sp.
XX
XX WO9858962-A1.
FN
PD 30-DEC-1998.
XX
PF 22-JUN-1998; 98WO-JP002765.
XX
PR 23-JUN-1997; 97JP-00165437.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Hinuma S, Kawamata Y, Fujii R, Matsumoto H;
XX WPI; 1999-105614/09.
XX
PS Use of G protein-coupled receptor ligands - for modulating prolactin
XX secretion or placental function, e.g. for treating menopausal syndrome,
XX tumours, autoimmune disease or abnormal pregnancy.
XX
PS Claim 3; Page 136; 24pp; English.
XX
CC The present sequence represents a bovine pituitary-derived ligand
XX fragment. It is used in the course of the invention. The specification
XX describes an agent for modulating prolactin secretion which comprises a
XX ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)
XX protein. The agents for promoting prolactin secretion can be used for
XX treating or preventing hypovarianism, gonocyst cacogenesis, menopausal
XX syndrome, euthyroid or hypometabolism. They can be used for promoting
XX lactation in a domestic mammal and as an aphrodisiac. The agents for
XX inhibiting prolactin secretion can be used for treating or preventing
XX pituitary adenomatosis, brain tumour, emmenopathy, autoimmune disease,
XX prolactinoma, infertility, impotence, amenorrhea, galactorrhea,
XX acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome, Forbes-
XX Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia. The
XX inhibitory agents can also be used as contraceptives. The agents for
XX modulating placental function can be used for treating or preventing
XX choriocarcinoma, hydatid mole, interruption mole, abortion, unthrifty fetus,
XX abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 78; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRG 13
| | | | | | | | | | | | | | |
Db 1 TPDINPAWYAGRG 13

RESULT 7
AAV49301
ID AAV49301 standard; peptide; 20 AA.
XX
AC AAV49301;
XX
DT 22-FEB-2000 (first entry)
XX
DE 19P2 ligand peptide fragment.
XX
KW Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;
KW pituitary; regulatory mechanism; central nervous system; pancreatic.
XX
OS Bos sp.
XX
XX Key Location/Qualifiers
XX Modified-site 20 /note= "C-terminal amide"
XX
XX WO9960112-A1.
FN
XX
PD 25-NOV-1999.
XX
PF 20-MAY-1999; 99WO-JP002650.
XX
PR 21-MAY-1998; 98JP-00140293.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Matsumoto H, Kitada C, Hinuma S;
XX WPI; 2000-039381/03.
XX
XX New monoclonal antibodies, useful in diagnosis, as drugs and in studying
XX diseases related to ligand abnormality.
XX
XX Disclosure; Page 27; 73pp; Japanese.
XX
CC The invention provides a monoclonal antibody which has a specific
XX reaction with the part peptide of the C-terminal of 19P2 ligand or its
XX derivative. The antibodies can be used in diagnosis or to treat or
XX prevent diseases associated with abnormality in the pituitary function
XX regulatory mechanism (e.g. promotion of prolactin secretion), central
XX nervous regulatory mechanism, and pancreatic function regulatory
XX mechanism. The antibody-based immunoassay can also be applied in
XX clarifying the physiological functions of the ligand and its derivative.
XX Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 78; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRG 13
| | | | | | | | | | | | | | |
Db 1 TPDINPAWYAGRG 13

RESULT 8
AAB10350
ID AAB10350 standard; peptide; 20 AA.
XX
AC AAB10350;
XX
DT 24-NOV-2000 (first entry)


```

XX DE Bovine oxytocin secretion promoting peptide SEQ ID NO: 6.
XX KW Bovine; oxytocin secretion promoter; G protein-coupled receptor protein;
XX KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
XX KW caesarean section; artificial fertilization; galactostasis; goat; pig;
XX KW veterinary medicine; milk production.
XX OS Bos taurus.
XX PN WO200038704-A1.
XX PD 06-JUL-2000.
XX PF 22-DEC-1999; 99WO-JP007199.
XX PR 25-DEC-1998; 98JP-00369585.
XX PA (TAKE ) TAKEDA CHEM IND LTD.
XX PI Matsumoto H, Kitada C, Hinuma S;
XX DR WPI; 2000-452298/39.
XX KW Physiologically-active polypeptide recognized as ligand by G protein-
XX KW coupled receptor protein, for promoting secretion of oxytocin, as drugs
XX KW for diseases relating to oxytocin secretion and in veterinary medicine.
XX PS Claim 5; Page 51; 72pp; Japanese.
XX CC This invention describes a novel oxytocin secretion-regulating agent
XX CC which contains a ligand peptide or its salt for the G protein-coupled
XX CC receptor protein. It is useful in the form of drugs for ameliorating,
XX CC preventing and treating diseases relating to oxytocin secretion e.g. weak
XX CC pains and atonic bleeding, before and after expulsion of placenta,
XX CC uterine recovery failure, caesarean section, stoppage of artificial
XX CC fertilization or galactostasis and is also applicable in veterinary
XX CC medicine for promoting milk production in cow, goat and pig. This
XX CC sequence represents a bovine peptide which acts as an oxytocin secretion
XX CC promoter
XX SQ Sequence 20 AA;
    Query Match 100.0%; Score 78; DB 3; Length 20;
    Best Local Similarity 100.0%; Pred. No. 2.8e-06;
    Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRG 13
    |||||
Db 1 TPDINPAWYAGRG 13

RESULT 9
AAB90996
XX ID AAB90996 standard; peptide; 20 AA.
XX AC AAB90996;
XX KW Prolactin releasing peptide SEQ ID NO:170.
XX DT 22-JUN-2001 (first entry)
XX DE Prolactin releasing peptide SEQ ID NO:170.
XX KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
XX KW blood component; modification; succinimide; maleimide group; amino;
XX KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN WO200069900-A2.
XX PD 23-NOV-2000.
XX PR 03-AUG-2000; 2000WO-FI000664.
XX PF 03-AUG-1999; 99US-00365756.
XX PR 03-AUG-2000; 2000WO-US013576.
XX PR 17-MAY-1999; 99US-0134406P.
XX PR 10-SEP-1999; 99US-0153406P.
XX PR 15-OCT-1999; 99US-0159783P.
XX PA (CONJ-) CONJUCHEM INC.
XX PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
XX DR WPI; 2001-112059/12.
XX KW Modifying and attaching therapeutic peptides to albumin prevents
XX KW peptidase degradation, useful for increasing length of in vivo activity.
XX PS Disclosure; Page 245; 733pp; English.
XX CC The present invention describes a modified therapeutic peptide (I)
XX CC comprising a therapeutically active amino acid region (III) and a
XX CC reactive group (II) (e.g. succinimide) and maleimide groups) attached to
XX CC a less therapeutically active amino acid region (IV), which covalently
XX CC bonds with amino/hydroxyl/thiol groups on blood components to form a
XX CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
XX CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
XX CC factors and neurotransmitters, to protect them from peptidase activity in
XX CC vivo for the treatment of various disorders. Endogenous therapeutic
XX CC peptides are not suitable as drug candidates as they require frequent
XX CC administration due to rapid degradation by peptidases in the body.
XX CC Modifying and attaching therapeutic peptides to albumin prevents or
XX CC reduces the action of peptidases to increase length of activity (half
XX CC life) and specificity as bonding to large molecules decreases
XX CC intracellular uptake and interference with physiological processes.
XX CC AAB90829 to AAB92441 represent peptides which can be used in the
XX CC exemplification of the present invention
XX SQ Sequence 20 AA;
    Query Match 100.0%; Score 78; DB 4; Length 20;
    Best Local Similarity 100.0%; Pred. No. 2.8e-06;
    Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRG 13
    |||||
Db 1 TPDINPAWYAGRG 13

RESULT 10
AAB46954
XX ID AAB46954 standard; protein; 20 AA.
XX AC AAB46954;
XX DT 04-MAY-2001 (first entry)
XX DE Peptide PRP20 fragment.
XX KW GPR10; UHR-1; PRP receptor; prolactin-releasing peptide; pain;
XX KW central nervous system disorder; autonomic regulation; analgesic;
XX KW hypotensive; blood pressure.
XX OS Unidentified.
XX KW Key Location/Qualifiers
XX KW Modified-site 20
XX KW /note= "C-terminal amide"
XX PN WO200109182-A1.
XX PD 08-FEB-2001.
XX PF 03-AUG-2000; 2000WO-FI000664.
XX PR 03-AUG-1999; 99US-00365756.

```

PR 20-MAR-2000; 2000US-00531567.
 XX (JUVA-) JUVANTIA PHARMA LTD OY.
 XX Panula Paj, Pertovaara A, Kalso E, Korpi E;
 XX WPI; 2001-182941/18.
 XX C-terminal fragments of prolactin-releasing peptide useful for regulating
 PT autonomic functions and in the manufacture of a medicament for regulating
 PT blood pressure.
 XX
 XX Claim 1; Page 10; 40pp; English.
 XX
 XX This invention describes a novel C-terminal fragment (I) of an isolated
 CC prolactin-releasing peptide (PrRP), referred to as PrRP20 and having a
 CC sequence (S1). The invention also describes (1) a therapeutic composition
 CC (C1) comprising (I) or a C-terminal fragment of PrRP referred to as PrRP8
 CC and comprising a sequence (S2); (2) a diagnostic method based on antisera
 CC against PrRP20 for identification of disorders involving the central
 CC nervous system, including those associated with pain or autonomic
 CC regulation, where specific antisera against the N-and/or C-terminal
 CC domains of PrRP is used to identify alterations in PrRP synthesis or
 CC levels; (3) a rat or human receptor encoded by a 1122 nucleotide sequence
 CC (S3), fully defined in the specification; (4) treating a person suffering
 CC from a disorder regulated by a receptor (II) encoded by a sequence of
 CC the central nervous system, by administering an agonist or antagonist to
 CC the receptor; and (5) treating blood pressure, by blocking of receptors
 CC of PrRP or its C-terminal fragment GlyileArgProValGlyArgPhe-NH₂ (S2).
 CC The products of the invention have analgesic and hypotensive activity.
 CC (I) is useful for regulating autonomic functions, such as increasing
 CC blood pressure. (I) is useful for treating pain, for manufacturing a
 CC medicament for regulating blood pressure, and for treating pain. Agonist
 CC and antagonist of (II) are useful for treating acute pain, inflammatory
 CC pain and neuropathic pain, for regulating autonomic functions and
 CC treating high blood pressure
 XX
 XX Sequence 20 AA;

Query Match 100.0%; Score 78; DB 4; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.8e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWVAGRG 13
 |||||
 Db 1 TPDINPAWVAGRG 13

RESULT 11
 AAG62519
 ID AAG62519 standard; peptide; 20 AA.
 XX
 AC AAG62519;
 XX

DT 24-AUG-2001 (first entry)
 XX

DE Bovine CRH releasing protein related peptide SEQ ID NO: 6.

KW Cow; corticotrophin releasing hormone; CRH; G protein receptor ligand;
 KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;
 KW Addison's disease; adrenal gland hyperfunction; obesity.
 XX

OS Bos taurus.

XX WO200135984-A1.

XX 25-MAY-2001.

XX 17-NOV-2000; 2000WO-JP008119.
 XX

PR 18-NOV-1999; 99JP-00327900.
 PR 26-SEP-2000; 2000JP-00297073.

XX (TAKE) TAKEDA CHEM IND LTD.
 XX
 XX Kitada C, Matsumoto H, Hinuma S;
 XX WPI; 2001-355552/37.
 XX

PT Use of G protein receptor ligand or peptide for controlling corticotropin
 PT releasing hormone secretion.
 XX

PS Claim 4; Page 64; 90pp; Japanese.

XX The present sequence describes a method of controlling the secretion of
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein
 CC receptor ligand. This can be used to control the secretion of CRH and is
 CC useful as an analgesic or for treating, preventing or ameliorating
 CC diseases associated with CRH secretion such as hyperaldosteronism,
 CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The
 CC present sequence is a peptide used in the exemplification of the
 CC invention
 XX

SQ Sequence 20 AA;

Query Match 100.0%; Score 78; DB 4; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.8e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWVAGRG 13
 |||||
 Db 1 TPDINPAWVAGRG 13

RESULT 12
 AAE26402
 ID AAE26402 standard; peptide; 20 AA.
 XX

AC AAE26402;

XX 13-DEC-2002 (first entry)

DE Bovine PrRP-31 C-terminal peptide, PrRP-20.

XX Wakefulness; sleep disorder; prolactin releasing peptide receptor; PrRP;
 KW GPR10; therapy; epilepsy; narcolepsy; sleepiness; sleep apnoea; insomnia;
 KW idiopathic hypersomnia; psychogenic hypersomnia; seizure; anticonvulsant;
 KW bovine.
 XX

OS Bos taurus.

XX US2002037533-A1.

XX 28-MAR-2002.

XX 17-AUG-2001; 2001US-00932161.

XX 26-APR-2000; 2000US-00560915.

XX (CIVE/) CIVELLI O.

XX (LINS/) LIN S.

XX Civelli O, Lin S;

XX WPI; 2002-403931/43.

PT Screening for compounds useful for promoting wakefulness or sleep, and
 PT for treating sleeping disorders, e.g. insomnia, hypersomnia or sleep
 PT apnea, comprises administering a prolactin releasing peptide agonist or
 PT antagonist.
 XX

PS Disclosure; Page 24; 35pp; English.

CC The present invention relates to a method of screening for compounds for
 CC promoting wakefulness or sleep in a mammal. The method involves
 CC administering a prolactin releasing peptide (PRP) receptor (GPR10)
 CC agonist or antagonist respectively and determining the ability of the
 CC compound to promote wakefulness or sleep. The compounds identified from
 CC the method are used in the therapy of epilepsy and other diseases
 CC associated with absence seizures and in promoting wakefulness and sleep
 CC in individuals having sleep disorders such as insomnia and narcolepsy.
 CC PRP receptor agonists may be used to treat common disorders which lead
 CC to sleepiness, e.g. sleep apnoea, narcolepsy, idiopathic hypersomnia and
 CC psychogenic hypersomnia. PRP receptor antagonists are useful for
 CC promoting sleep and for treating insomnia such as adjustment sleep
 CC disorder and psychophysiological insomnia. The present sequence is bovine
 CC PRP-31 C-terminal peptide, P-RP-20
 XX
 XX Sequence 20 AA;

Query Match 100.0%; Score 78; DB 5; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.8e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRG 13
 |||||
 Db 1 TPDINPAWYAGRG 13

RESULT 13

ABU60834
 ID ABU60834 standard; peptide; 20 AA.

XX AC ABU60834;

DT 06-MAY-2003 (first entry)

DE Peptide production by gene recombination associated peptide #18.

XX Peptide production; low-molecular peptide; KiSS-1; GPR8 ligand;
 KW gene recombination.

XX Bos sp.

PN WO200292829-A1.

XX 21-NOV-2002.

PF 16-MAY-2002; 2002WO-JP004735.

PR 17-MAY-2001; 2001JP-00147341.

XX (TAKE) TAKEDA CHEM IND LTD.

PI Nishimura O, Suenaga M, Ito T, Kitada C;

XX WPI; 2003-129302/12.

PT Process for producing peptides e.g. KiSS-1 peptide and GPR8 ligand for
 PT subsequent applications by gene recombination technique through tandem
 PT repeats to provide precursor protein with specific cleavage sites.

XX Disclosure; Page 62; 87pp; Japanese.

XX The invention describes a method of producing a peptide comprising the
 CC excision of the N and C-terminals of a target peptide with enzymes or
 CC chemically through the attached cleavage sites repeated by ligation in a
 CC precursor protein. The method is for producing (low-molecular) peptides
 CC e.g. KiSS-1 peptide and GPR8 ligand for subsequent applications by the
 CC gene recombination technique through tandem repeats to provide a
 CC precursor protein with specific cleavage sites. With this method, peptide
 CC production can be carried out easily to provide large quantities of the
 CC required peptides. This is the amino acid sequence of a peptide
 CC associated with the peptide production method of the invention
 XX
 XX Sequence 20 AA;

Query Match 100.0%; Score 78; DB 6; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.8e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRG 13
 |||||
 Db 1 TPDINPAWYAGRG 13

RESULT 14

AAW31375
 ID AAW31375 standard; peptide; 21 AA.

XX AC AAW31375;

DT 06-APR-1998 (first entry)

DE Bovine G protein-coupled receptor ligand peptide fragment 5.

XX G protein-coupled receptor; ligand binding; pharmaceutical; modulator;
 KW pituitary; central nervous system; pancreas; prophylactic;
 XX therapeutic agent.

XX Bos taurus.

PN WO9724436-A2.

PD 10-JUL-1997.

PF 26-DEC-1996; 96WO-JP003821.

XX 28-DEC-1995; 95JP-00343371.

PR 15-MAR-1996; 96JP-00059419.

PR 12-AUG-1996; 96JP-00211805.

PR 18-SEP-1996; 96JP-00246573.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;

PI Kitada C;

XX WPI; 1997-363672/33.

XX N-PSDB; AAV02398.

XX Ligand peptide for G protein-coupled receptor - acts by modulating
 PT function in the central nervous system, pancreas and pituitary gland.

PS Claim 2; Page 162; 258pp; English.

XX This sequence represents a peptide fragment of a novel bovine pituitary
 CC derived ligand corresponding to amino acid residues 34 to 54 of the
 CC sequence in AAW31375 and is used in an assay to monitor ligand binding to
 CC the G protein-coupled receptor protein. Pharmaceutical compositions
 CC containing this ligand may be used as a pituitary function modulator, a
 CC central nervous system modulator or a pancreatic function modulator. This
 CC ligand could have specific applications as a prophylactic or therapeutic
 CC agent for dementia, depression, hyperkinetic syndrome, disturbance of
 CC consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone
 CC secretory disease, hyper- and polyphagia, hypercholesterolaemia,
 CC hyperglycaemia, hyperlipidaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia,
 CC amyotrophic lateral sclerosis, acute myocardial infarction,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis, asthma, epilepsy, infertility and/or oligogalactia. Assays
 CC can also be developed to screen compounds which are capable of altering
 CC the binding activity of the ligand thus affecting activation of the G
 CC protein-coupled receptor protein
 XX
 XX Sequence 21 AA;

Query Match 100.0%; Score 78; DB 2; Length 21;

Best Local Similarity 100.0%; Pred. No. 2.9e-06; Mismatches 0; Gaps 0;
Matches 13; Conservative 0; Indels 0; Gaps 0;
QY 1 TPDINPAWYAGRG 13
| | | | | | | | | | | | |
Db 1 TPDINPAWYAGRG 13
| | | | | | | | | | | | |
RESULT 15
AAW95192
ID AAW95192 standard; peptide; 21 AA.
XX
AC AAW95192;
XX
DT 10-MAR-1999 (first entry)
XX
DE Bovine pituitary-derived ligand polypeptide fragment.
XX
KW Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
KW Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;
KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
KW gene therapy; transgenic animal; bovine.
XX
OS Bos sp.
XX
PN WO9849295-A1.
XX
PD 05-NOV-1998.
XX
PF 27-APR-1998; 98WO-JP001923.
XX
PR 28-APR-1997; 97JP-00109974.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Hinuma S, Fukusumi S;
XX
DR WPI; 1999-009423/01.
XX
PT New polypeptide ligand for orphan G protein coupled receptors - used for
PT treating disorders of central nervous system, pituitary and pancreas, and
PT for drug screening.
XX
PS Example 19; Page 151; 206pp; English.
XX
CC The invention relates to a murine pituitary-derived ligand polypeptide
CC which is a ligand for the G-protein coupled orphan receptor designated
CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing
CC the ligand polypeptide encoding DNA are used to produce a recombinant
CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate
CC function of the pituitary, central nervous system, pancreas and other
CC tissues and can be used to screen for agents that modulate binding of the
CC polypeptide to the receptor; to quantify the amount of receptor in a
CC sample and to raise antibodies. They may also be used therapeutically,
CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's
CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;
CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;
CC rheumatoid arthritis, epilepsy and many others, also to improve post-
CC operative nutritional status and as vasopressor. Transgenic animals
CC carrying the ligand polypeptide encoding DNA or its mutin are used to
CC study the function of the polypeptide-expressing genes, as models of
CC disease, for drug screening and as source of cell lines. The ligand
CC polypeptide DNA is used as a source of probes and primers; to identify
CC related sequences; in receptor-binding assays; for production of Ab and
CC antisera; in drug development; for gene therapy and to develop transgenic
CC animals. The present sequence represents a bovine genome-derived ligand
CC polypeptide fragment which is similar to the murine ligand-polypeptide
XX Sequence 21 AA;
SQ

Query Match 100.0%; Score 78; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.9e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TPDINPAWYAGRG 13
| | | | | | | | | | | | |
Db 1 TPDINPAWYAGRG 13
| | | | | | | | | | | | |
Search completed: August 12, 2004, 14:43:54
Job time : 28.7122 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:49:10 ; Search time 22.9767 Seconds
(without alignments)
177.617 Million cell updates/sec

Title: US-09-700-643A-1_COPY_12_24

Perfect score: 78

Sequence: 1 TPDINPAWYAGRG 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1292805 seqs, 313927144 residues

Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	100.0	19	13	US-10-044-592-27
2	78	100.0	20	9	US-09-932-161-16
3	78	100.0	20	13	US-10-044-592-42
4	78	100.0	20	14	US-10-096-777-16
5	78	100.0	21	13	US-10-044-592-43
6	78	100.0	22	13	US-10-044-592-44
7	78	100.0	25	13	US-10-044-592-78
8	78	100.0	29	13	US-10-044-592-26
9	78	100.0	31	9	US-09-932-161-13
10	78	100.0	31	13	US-10-044-592-39
11	78	100.0	31	14	US-10-096-777-13
12	78	100.0	32	13	US-10-044-592-40
13	78	100.0	33	13	US-10-044-592-41
14	78	100.0	98	13	US-10-044-592-28
15	78	100.0	98	13	US-10-044-592-38

16	78	100.0	98	13	US-10-044-592-82	Sequence 82, Appl
17	78	100.0	98	13	US-10-044-592-84	Sequence 84, Appl
18	78	100.0	98	13	US-10-044-592-86	Sequence 86, Appl
19	78	100.0	98	13	US-10-044-592-88	Sequence 88, Appl
20	74	94.9	20	9	US-09-932-161-17	Sequence 17, Appl
21	74	94.9	20	13	US-10-044-592-6	Sequence 6, Appl
22	74	94.9	20	14	US-10-096-777-17	Sequence 17, Appl
23	74	94.9	31	9	US-09-932-161-14	Sequence 14, Appl
24	74	94.9	31	13	US-10-044-592-4	Sequence 4, Appl
25	74	94.9	31	13	US-10-044-592-5	Sequence 5, Appl
26	74	94.9	31	14	US-10-096-777-14	Sequence 14, Appl
27	74	94.9	70	13	US-10-044-592-90	Sequence 90, Appl
28	74	94.9	82	13	US-10-044-592-1	Sequence 1, Appl
29	74	94.9	86	13	US-10-044-592-96	Sequence 96, Appl
30	74	94.9	91	13	US-10-044-592-94	Sequence 94, Appl
31	72	92.3	20	9	US-09-932-161-18	Sequence 18, Appl
32	72	92.3	20	14	US-10-096-777-18	Sequence 18, Appl
33	72	92.3	31	9	US-09-932-161-15	Sequence 15, Appl
34	72	92.3	31	14	US-10-096-777-15	Sequence 15, Appl
35	72	92.3	87	13	US-10-044-592-92	Sequence 92, Appl
36	57	73.1	9	13	US-10-044-592-8	Sequence 8, Appl
37	53	67.9	465	14	US-10-301-822-197	Sequence 197, App
38	50	64.1	527	15	US-10-369-493-5327	Sequence 5327, Ap
39	49	62.8	183	12	US-10-424-599-268092	Sequence 268092,
40	44	56.4	109	12	US-10-424-599-218693	Sequence 218693,
41	44	56.4	398	15	US-10-225-068-74	Sequence 74, Appl
42	44	56.4	398	15	US-10-374-780A-356	Sequence 356, App
43	44	56.4	816	16	US-10-437-963-198301	Sequence 198301,
44	43	55.1	387	12	US-10-282-122A-50882	Sequence 50882, A
45	43	55.1	548	15	US-10-374-750A-2170	Sequence 2170, Ap

ALIGNMENTS

RESULT 1
US-10-044-592-27
; Sequence 27, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 27
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Bovine
; US-10-044-592-27

Query Match 100.0%; Score 78; DB 13; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRG 13
Db 1 TPDINPAWYAGRG 13

RESULT 2
US-09-932-161-16
; Sequence 16, Application US/09932161
; Patent No. US20020037533A1

```

; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-932-161-16

```

```

Query Match      100.0%; Score 78; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TPDINPAWYAGRG 13
DB      1 TPDINPAWYAGRG 13

```

RESULT 3

```

US-10-044-592-42
; Sequence 42, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 42
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-42

```

```

Query Match      100.0%; Score 78; DB 13; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TPDINPAWYAGRG 13
DB      1 TPDINPAWYAGRG 13

```

RESULT 4

```

US-10-096-777-16
; Sequence 16, Application US/10096777
; Publication No. US20030171270A1
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777
; CURRENT FILING DATE: 2002-03-12

```

```

; PRIOR APPLICATION NUMBER: US/09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Bos taurus
US-10-096-777-16

```

```

Query Match      100.0%; Score 78; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TPDINPAWYAGRG 13
DB      1 TPDINPAWYAGRG 13

```

RESULT 5

```

US-10-044-592-43
; Sequence 43, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 43
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-43

```

```

Query Match      100.0%; Score 78; DB 13; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TPDINPAWYAGRG 13
DB      1 TPDINPAWYAGRG 13

```

RESULT 6

```

US-10-044-592-44
; Sequence 44, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 44

```

```
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-44

Query Match      100.0%; Score 78; DB 13; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRG 13
   |||||
Db 1 TPDINPAWYAGRG 13

RESULT 7
US-10-044-592-78
; Sequence 78, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 78
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Bovine
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(20)
; OTHER INFORMATION: primer
; NAME/KEY: misc.feature
; LOCATION: (52)..(76)
; OTHER INFORMATION: primer
US-10-044-592-78

Query Match      100.0%; Score 78; DB 13; Length 25;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRG 13
   |||||
Db 10 TPDINPAWYAGRG 22

RESULT 8
US-10-044-592-26
; Sequence 26, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96

Query Match      100.0%; Score 78; DB 13; Length 31;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRG 13
   |||||
Db 12 TPDINPAWYAGRG 24

RESULT 9
US-09-932-161-13
; Sequence 13, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Civeilli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-932-161-13

Query Match      100.0%; Score 78; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRG 13
   |||||
Db 12 TPDINPAWYAGRG 24

RESULT 10
US-10-044-592-39
; Sequence 39, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 39
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-39

Query Match      100.0%; Score 78; DB 13; Length 31;
```

Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRG 13
Db 12 TPDINPAWYAGRG 24

RESULT 11
US-10-096-777-13
; Sequence 13, Application US/10096777
; Publication No. US20030171270A1
; GENERAL INFORMATION:
; APPLICANT: Civeilli, Olivier
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/560,915
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Bos taurus
US-10-096-777-13

Query Match 100.0%; Score 78; DB 14; Length 31;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRG 13
Db 12 TPDINPAWYAGRG 24

RESULT 12
US-10-044-592-40
; Sequence 40, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 40
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-40

Query Match 100.0%; Score 78; DB 13; Length 32;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRG 13
Db 12 TPDINPAWYAGRG 24

RESULT 13
US-10-044-592-41
; Sequence 41, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 41
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-41

Query Match 100.0%; Score 78; DB 13; Length 33;
Best Local Similarity 100.0%; Pred. No. 3.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRG 13
Db 12 TPDINPAWYAGRG 24

RESULT 14
US-10-044-592-28
; Sequence 28, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 28
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Murine
US-10-044-592-28

Query Match 100.0%; Score 78; DB 13; Length 98;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRG 13
Db 34 TPDINPAWYAGRG 46

RESULT 15
US-10-044-592-38
; Sequence 38, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:

; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 243US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 38
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-38

Query Match 100.0%; Score 78; DB 13; Length 98;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRG 13
| | | | | | | | | | | | | | | |
Db 34 TPDINPAWYAGRG 46

Search completed: August 12, 2004, 15:22:50
Job time : 23.9767 secs

Thu Aug 12 15:25:12 2004

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:36 ; Search time 7.40698 Seconds
(without alignments)
90.609 Million cell updates/sec

Title: US-09-700-643A-1_COPY_12_24

Perfect score: 78

Sequence: 1 TPDINPAMYAGRG 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/aaa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/aaa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/aaa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/aaa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/aaa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/2/aaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	78	100.0	19	3 US-09-105-678A-30	Sequence 30, Appl
2	78	100.0	19	3 US-08-776-971-4	Sequence 4, Appl
3	78	100.0	19	3 US-09-421-208-30	Sequence 30, Appl
4	78	100.0	20	3 US-09-105-678A-34	Sequence 34, Appl
5	78	100.0	20	3 US-08-776-971-8	Sequence 8, Appl
6	78	100.0	20	3 US-08-776-971-98	Sequence 98, Appl
7	78	100.0	20	3 US-09-421-208-34	Sequence 34, Appl
8	78	100.0	20	4 US-09-560-915-16	Sequence 16, Appl
9	78	100.0	21	3 US-09-105-678A-35	Sequence 35, Appl
10	78	100.0	21	3 US-08-776-971-9	Sequence 9, Appl
11	78	100.0	21	3 US-09-421-208-35	Sequence 35, Appl
12	78	100.0	22	3 US-09-105-678A-36	Sequence 36, Appl
13	78	100.0	22	3 US-08-776-971-10	Sequence 10, Appl
14	78	100.0	22	3 US-09-421-208-36	Sequence 36, Appl
15	78	100.0	25	3 US-08-776-971-111	Sequence 111, Appl
16	78	100.0	29	3 US-09-105-678A-29	Sequence 29, Appl
17	78	100.0	29	3 US-08-776-971-3	Sequence 3, Appl
18	78	100.0	29	3 US-09-421-208-29	Sequence 29, Appl
19	78	100.0	31	3 US-09-105-678A-7	Sequence 7, Appl
20	78	100.0	31	3 US-09-105-678A-31	Sequence 31, Appl
21	78	100.0	31	3 US-08-776-971-5	Sequence 5, Appl
22	78	100.0	31	3 US-08-776-971-97	Sequence 97, Appl
23	78	100.0	31	3 US-09-421-208-7	Sequence 7, Appl
24	78	100.0	31	3 US-09-421-208-31	Sequence 31, Appl
25	78	100.0	31	4 US-09-560-915-13	Sequence 13, Appl
26	78	100.0	32	3 US-09-105-678A-32	Sequence 32, Appl
27	78	100.0	32	3 US-08-776-971-6	Sequence 6, Appl

28	78	100.0	32	3 US-09-421-208-32	Sequence 32, Appl
29	78	100.0	33	3 US-09-105-678A-33	Sequence 33, Appl
30	78	100.0	33	3 US-08-776-971-7	Sequence 7, Appl
31	78	100.0	33	3 US-09-421-208-33	Sequence 33, Appl
32	78	100.0	98	3 US-08-776-971-1	Sequence 1, Appl
33	78	100.0	98	3 US-08-776-971-44	Sequence 44, Appl
34	78	100.0	98	3 US-08-776-971-122	Sequence 122, Appl
35	78	100.0	98	3 US-08-776-971-131	Sequence 131, Appl
36	78	100.0	98	3 US-08-776-971-136	Sequence 136, Appl
37	74	94.9	20	3 US-09-105-678A-40	Sequence 40, Appl
38	74	94.9	20	3 US-08-776-971-50	Sequence 50, Appl
39	74	94.9	20	3 US-09-421-208-40	Sequence 40, Appl
40	74	94.9	20	4 US-09-560-915-17	Sequence 17, Appl
41	74	94.9	21	3 US-09-105-678A-41	Sequence 41, Appl
42	74	94.9	21	3 US-08-776-971-51	Sequence 51, Appl
43	74	94.9	21	3 US-09-421-208-41	Sequence 41, Appl
44	74	94.9	22	3 US-09-105-678A-42	Sequence 42, Appl
45	74	94.9	22	3 US-08-776-971-52	Sequence 52, Appl

ALIGNMENTS

RESULT 1
US-09-105-678A-30
; Sequence 30, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-105-678A-30

Query Match 100.0%; Score 78; DB 3; Length 19;
Best Local Similarity 100.0%; Pred. No. 8e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAMYAGRG 13

Db 1 TPDINPAWYAGRG 13
|||||
RESULT 2
US-08-776-971-4
; Sequence 4, Application US/08776971B
; Patent No. 5228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; Habata, Yugo
; Kawamata, Yuji
; Hosoya, Masaki
; Fujii, Ryo
; Fukusumi, Shoji
; Kitada, Chieko
; TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-FEB-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-776-971-4
Query Match 100.0%; Score 78; DB 3; Length 19;
Best Local Similarity 100.0%; Pred. No. 8e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TPDINPAWYAGRG 13
Db 1 TPDINPAWYAGRG 13
|||||
RESULT 3
US-09-700-643a-1
; Sequence 30, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-421-208-30
Query Match 100.0%; Score 78; DB 3; Length 19;
Best Local Similarity 100.0%; Pred. No. 8e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TPDINPAWYAGRG 13
Db 1 TPDINPAWYAGRG 13
|||||
RESULT 4
US-09-105-678A-34
; Sequence 34, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-34

Query Match 100.0%; Score 78; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.4e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRG 13
Db 1 TPDINPAWYAGRG 13

RESULT 5
US-08-776-971-8
; Sequence 8, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; Habata, Yugo
; Kawamata, Yuji
; Hosoya, Masaki
; Fujii, Ryo
; Fukusumi, Shoji
; Kitada, Chieko
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,971B
FILING DATE: 06-Feb-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03821
FILING DATE: 28-DEC-1996
APPLICATION NUMBER: JP 7/343371
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: JP 8/59419
FILING DATE: 15-MAR-1996

APPLICATION NUMBER: JP 8/211805
FILING DATE: 12-AUG-1996
APPLICATION NUMBER: JP 8/246573
FILING DATE: 18-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 47176
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-08-776-971-8

Query Match 100.0%; Score 78; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.4e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRG 13
Db 1 TPDINPAWYAGRG 13

RESULT 6
US-08-776-971-98
; Sequence 98, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; Habata, Yugo
; Kawamata, Yuji
; Hosoya, Masaki
; Fujii, Ryo
; Fukusumi, Shoji
; Kitada, Chieko
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,971B
FILING DATE: 06-Feb-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03821
FILING DATE: 28-DEC-1996
APPLICATION NUMBER: JP 7/343371
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: JP 8/59419
FILING DATE: 15-MAR-1996
APPLICATION NUMBER: JP 8/211805
FILING DATE: 12-AUG-1996
APPLICATION NUMBER: JP 8/246573
FILING DATE: 18-SEP-1996
ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 47176
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 98:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 98:

Query Match 100.0%; Score 78; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.4e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRG 13
Db 1 TPDINPAWYAGRG 13

RESULT 7

US-09-421-208-34
Sequence 34, Application US/09421208
Patent No. 6258561

GENERAL INFORMATION:
APPLICANT: Suenaga, Masato

APPLICANT: Moriya, Takeo

APPLICANT: Tanaka, Yoko

APPLICANT: Nishimura, Osamu

TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street
CITY: Boston

STATE: MA
COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/421,208

FILING DATE:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/105,678

FILING DATE: 26-JUN-1998

APPLICATION NUMBER: JP 172118/1997

FILING DATE: 27-JUN-1997

ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 48466-342

TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids
TYPE: amino acid

STRANDEDNESS:
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-09-421-208-34

Query Match 100.0%; Score 78; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.4e-07;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRG 13

Db 1 TPDINPAWYAGRG 13

RESULT 8

US-09-560-915-16

Sequence 16, Application US/09560915

Patent No. 6383764

GENERAL INFORMATION:
APPLICANT: Civeilli, Olivier

APPLICANT: Lin, Steven

TITLE OF INVENTION: Therapeutic Compositions and Methods

FILE REFERENCE: P-UC 3534

CURRENT APPLICATION NUMBER: US/09/560,915

CURRENT FILING DATE: 2000-04-28

NUMBER OF SEQ ID NOS: 24

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 16

LENGTH: 20

TYPE: PRT

ORGANISM: Bos taurus

US-09-560-915-16

Query Match 100.0%; Score 78; DB 4; Length 20;

Best Local Similarity 100.0%; Pred. No. 8.4e-07;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRG 13

Db 1 TPDINPAWYAGRG 13

RESULT 9

US-09-105-678A-35

Sequence 35, Application US/09105678A

Patent No. 6103882

GENERAL INFORMATION:
APPLICANT: Suenaga, Masato

APPLICANT: Moriya, Takeo

APPLICANT: Tanaka, Yoko

APPLICANT: Nishimura, Osamu

TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street
CITY: Boston

STATE: MA
COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A

FILING DATE: 26-JUN-1998

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997

FILING DATE: 27-JUN-1997

ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 48466-342

TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 35:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-105-678A-35

Query Match 100.0%; Score 78; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 8.9e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRG 13
DB 1 TPDINPAWYAGRG 13

RESULT 10
US-08-776-971-9
; Sequence 9, Application US/08776971B
; Patent No. 6228994
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; HABATA, YUGO
; KAWAMATA, YUJI
; HOSOKAWA, MASAKI
; FUJII, RYO
; FUKUSUMI, SHOJI
; KITADA, CHIEKO
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JF96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal

; SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-08-776-971-9
Query Match 100.0%; Score 78; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 8.9e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRG 13
DB 1 TPDINPAWYAGRG 13

RESULT 11
US-09-421-208-35
; Sequence 35, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-421-208-35

Query Match 100.0%; Score 78; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 8.9e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRG 13
DB 1 TPDINPAWYAGRG 13

RESULT 12
US-09-105-678A-36
; Sequence 36, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:

```

```

; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-105-678A-36

```

```

Query Match 100.0%; Score 78; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 9.4e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TPDINPAWYAGRG 13
Db 1 TPDINPAWYAGRG 13

```

```

RESULT 13
US-08-776-971-10
; Sequence 10, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; Habata, Yugo
; Kawanata, Yuji
; Hosoya, Masaki
; Fujii, Ryo
; Fukusumi, Shoji
; Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible

```

```

; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-08-776-971-10

```

```

Query Match 100.0%; Score 78; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 9.4e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TPDINPAWYAGRG 13
Db 1 TPDINPAWYAGRG 13

```

```

RESULT 14
US-09-421-208-36
; Sequence 36, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998

```


APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-3440
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-421-208-36

Query Match 100.0%; Score 78; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 9.4e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRG 13
Db 1 TPDINPAWYAGRG 13

RESULT 15

US-08-776-971-111
Sequence 111, Application US/08776971B
Patent No. 6228984

GENERAL INFORMATION:

APPLICANT: Hinuma, Shuji
Kabata, Yugo
Kawamata, Yuji
Hosoya, Masaki
Fujii, Ryo
Fukusumi, Shoji
Kitada, Chieko

TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE

NUMBER OF SEQUENCES: 140

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/776,971B

FILING DATE: 06-Feb-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP96/03821

FILING DATE: 28-DEC-1996

APPLICATION NUMBER: JP 7/343371

FILING DATE: 28-DEC-1995

APPLICATION NUMBER: JP 8/59419

FILING DATE: 15-MAR-1996

APPLICATION NUMBER: JP 8/211805

FILING DATE: 12-AUG-1996

APPLICATION NUMBER: JP 8/246573

FILING DATE: 18-SEP-1996

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 47176

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 111:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 111:
US-08-776-971-111

Query Match 100.0%; Score 78; DB 3; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYAGRG 13
Db 10 TPDINPAWYAGRG 22

Search completed: August 12, 2004, 14:52:13
Job time : 7.40698 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:36 ; Search time 5.74419 Seconds
(without alignments)
217.697 Million cell updates/sec

Title: US-09-700-643a-2_COPY_12_24

Perfect score: 76
Sequence: 1 TPDINPAWYASRG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: Pirl:*
2: Pirl2:*
3: Pirl3:*
4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	89.5	83	JC7607	prolactin-releasin
2	45	59.2	767	T21969	hypothetical prote
3	44	57.9	527	T33175	hypothetical prote
4	43	56.6	1359	T10235	xanthine dehydrog
5	43	56.6	1364	T10236	xanthine dehydrog
6	42.5	55.9	664	P83376	conserved hypothet
7	42.5	55.9	779	T49717	related to BCS1 pr
8	42	55.3	221	A75366	probable transcrip
9	42	55.3	311	S66500	cytochrome-c oxida
10	42	55.3	364	B64766	yaiW protein - Esc
11	42	55.3	364	D90682	hypothetical prote
12	42	55.3	364	H85532	hypothetical prote
13	42	55.3	364	AH0548	probable lipoprote
14	42	55.3	367	G83309	conserved hypothet
15	42	55.3	476	G64720	probable amino aci
16	42	55.3	476	G90629	probable inner mem
17	42	55.3	476	G85480	inner membrane tra
18	42	55.3	512	T40815	probable ubiquitin
19	42	55.3	1296	T16959	hypothetical prote
20	41	53.9	132	B82824	protein-export mem
21	41	53.9	330	G82844	cysteine synthase
22	41	53.9	430	B69009	conserved hypothet
23	41	53.9	986	OYURGA	speract receptor p
24	40	52.6	176	T28762	hypothetical prote
25	40	52.6	181	S44765	C29E4.9 protein -
26	40	52.6	376	A48197	opsin, ocellar - A
27	40	52.6	376	B48197	opsin, lateral eye
28	40	52.6	476	AG0502	probable amino-aci
29	40	52.6	957	A80509	hypothetical prote

30	40	52.6	1711	2	T21432	hypothetical prote
31	40	52.6	1762	2	T03222	probable polyketid
32	40	52.6	2100	2	T03223	probable polyketid
33	40	52.6	7962	2	T38346	elastic titin, hu
34	39	51.3	163	2	G75424	bacterioferritin c
35	39	51.3	468	2	C83160	nitrite extrusion c
36	39	51.3	508	2	T05156	probable glucose t
37	39	51.3	612	2	E81287	probable sugar tra
38	39	51.3	822	2	E75523	ATP-dependent heli
39	38.5	50.7	240	2	D64688	probable 1-acylgly
40	38.5	50.7	352	2	T05418	hypothetical prote
41	38	50.0	105	2	A97505	hypothetical prote
42	38	50.0	105	2	AE2723	hypothetical prote
43	38	50.0	118	2	B72686	hypothetical prote
44	38	50.0	212	2	AH1944	phosphoglycerate m
45	38	50.0	282	2	T35304	probable peptidase

ALIGNMENTS

RESULT 1

JC7607
prolactin-releasing peptide - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C:Accession: JC7607
R:Yanada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Hosoya, T.
Biochem. Biophys. Res. Commun. 281, 53-56, 2001
A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene: Mu1
A:Reference number: JC7607; MUID:21092785; PMID:11178959
A:Contents: Spleen
A:Accession: JC7607
A:Molecule type: DNA
A:Residues: 1-83 <YAM>
A:Cross-references: DDBJ:AB040612; DDBJ:AB040613
C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior p
release, and stimulation of ACTH secretion from the pituitary.
C:Genetics:
A:Gene: PRP
A:Introns: 33/1

Query Match	89.5%	Score 68;	DB 2;	Length 83;
Best Local Similarity	84.6%	Pred. No. 0.00022;		
Matches	11;	Conservative	0;	Mismatches 2;
			Indels	0;
			Gaps	0;
Qy	1	TPDINPAWYASRG	13	
Db	33	TPDINPAWYTGRC	45	

RESULT 2

T21969
hypothetical protein F38E11.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21969
R:Mathews, P
submitted to the EMBL Data Library, January 1996
A:Reference number: Z19495
A:Accession: T21969
A:Status: preliminary; translated from GS/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-767 <WIL>
A:Cross-references: EMBL:Z68342; PIDN:CAA92775.1; GSPDB:GNO0022; CESP:F38E11.7
A:Experimental source: clone F38E11
C:Genetics:
A:Gene: CESP:F38E11.7
A:Map position: 4
A:Introns: 50/2; 118/1; 139/2; 189/3; 226/1; 248/1; 287/2; 375/2; 432/3; 465/3; 548/1;
Query Match 59.2%; Score 45; DB 2; Length 767;
Best Local Similarity 75.0%; Pred. No. 16;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PDINPAWY 9
Db 747 PDVKAQWY 754

RESULT 3

T33175
hypothetical protein C24G6.6 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C/Accession: T33175
R/Graco, T.; Bradshaw, H.; Keppler, D.
submitted to the EMBL Data Library, May 1998
A/Description: The sequence of C. elegans cosmid C24G6.
A/Reference number: Z21298
A/Accession: T33175
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-527 <GRE>
A/Cross-references: EMBL:AF067936; PIDN:AAC19213.1; GSPDB:GN00023; CESP:C24G6.6
A/Experimental source: strain Bristol N2; clone C24G6
C/Genetics:
A/Map position: 5
A/Introns: 20/3; 77/1; 129/2; 208/3; 470/2

Query Match 57.9%; Score 44; DB 2; Length 527;
Best Local Similarity 58.3%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PDINPAWYASRG 13
Db 370 PNVLSAWYAGRG 381

RESULT 4

T10235
xanthine dehydrogenase homolog T1111.130 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cross)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jan-2000
C/Accession: T10235
R/Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Jesse, T.; Heijnen, L.; Vos, P.; Mewes,
submitted to the Protein Sequence Database, June 1999
A/Reference number: Z16992
A/Accession: T10235
A/Molecule type: DNA
A/Residues: 1-1359 <BEV>
A/Cross-references: EMBL:AL079347; GSPDB:GN00062; ATSP:T1111.130
A/Experimental source: cultivar Columbia; BAC clone Fil111
C/Genetics:
A/Map position: 4
A/Introns: 43/1; 74/2; 110/3; 608/3; 689/3; 794/3; 874/3; 969/3; 1043/3; 1097/3; 1142/3;
C/Superfamily: xanthine dehydrogenase; ferredoxin [2Fe-2S] homology
C/Keywords: 2Fe-2S; metalloprotein
F/37-82/Domain: ferredoxin [2Fe-2S] homology <PDX>
F/51,56,59,81/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 56.6%; Score 43; DB 2; Length 1359;
Best Local Similarity 70.0%; Pred. No. 61;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 DINPAWYASR 12
Db 386 DLNPLWASR 395

RESULT 5

T10236
xanthine dehydrogenase homolog T1111.140 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cross)

C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 07-Dec-1999
C/Accession: T10236
R/Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Jesse, T.; Heijnen, L.; Vos, P.; Mewes,
submitted to the Protein Sequence Database, June 1999
A/Reference number: Z16992
A/Accession: T10236
A/Molecule type: DNA
A/Residues: 1-1364 <BEV>
A/Cross-references: EMBL:AL079347; GSPDB:GN00062; ATSP:T1111.140
A/Experimental source: cultivar Columbia; BAC clone Fil111
C/Genetics:
A/Map position: 4
A/Introns: 35/1; 66/2; 115/3; 613/3; 694/3; 799/3; 879/3; 974/3; 1048/3; 1102/3; 1147/3;
C/Superfamily: xanthine dehydrogenase; ferredoxin [2Fe-2S] homology
C/Keywords: 2Fe-2S; metalloprotein
F/43,48,51,86/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 55.6%; Score 43; DB 2; Length 1364;
Best Local Similarity 70.0%; Pred. No. 61;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 DINPAWYASR 12
Db 391 DLNPLWASR 400

RESULT 6

F83376
conserved hypothetical protein PA2151 [imported] - Pseudomonas aeruginosa (strain PA01)
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C/Accession: F83376
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 999-984, 2000
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A/Reference number: A82950; MUID:20437337; PMID:10984043
A/Accession: F83376
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-664 <STO>
A/Cross-references: GB:AE004642; GB:AE004091; NID:g9948163; PIDN:AA05539.1; GSPDB:GN001
A/Experimental source: strain PA01
C/Genetics:
A/Map position: 4
A/Introns: 1-664 <STO>

Query Match 55.9%; Score 42.5; DB 2; Length 664;
Best Local Similarity 66.7%; Pred. No. 35;
Matches 8; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

Qy 1 TPDINPAWYASR 12
Db 478 TPDINP-WFLQR 488

RESULT 7

T49717
related to BCS1 protein precursor [imported] - Neurospora crassa
N/Alternate names: protein B23L21.300
C/Species: Neurospora crassa
C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C/Accession: T49717
R/Schulte, U.; Aign, V.; Hobeisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A/Reference number: Z25022
A/Accession: T49717
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-779 <SCH>
A/Cross-references: EMBL:AL356172; GSPDB:GN00116; NCSP:B23L21.300
A/Experimental source: BAC clone B23L21; strain OR74A

C:Genetics:
A:Gene: NCSP:B23121.300
A:Map position: 6

Query Match 55.9%; Score 42.5; DB 2; Length 779;
Best Local Similarity 56.2%; Pred. No. 41;
Matches 9; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

QY 1 TPDINFA---WYASRG 13
| : ||| ||| : |||
Db 286 TDYLNPAATRRWYANRG 301

RESULT 8
A75366
probable transcription regulator - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: A75366
R.; White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, I.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: A75366
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-221 <WHI>
A:Cross-references: GB:AE002010; GB:AE000513; NID:G6459448; PIDN:AAF11240.1; PID:G645945
A:Experimental_source: strain R1
C:Genetics:
A:Gene: D31683
A:Map position: 1
C:Superfamily: Bacillus subtilis transcription regulator pairB

Query Match 55.3%; Score 42; DB 2; Length 221;
Best Local Similarity 87.5%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 INPAWVAS 11
| : ||| |||
Db 105 IDPAWVAS 112

RESULT 9
S66600
cytochrome-c oxidase (EC 1.9.3.1) chain III - blue mussel mitochondrion
C:Species: Mitochondrion Mytilus edulis (blue mussel)
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999
C:Accession: S66600; S28753
R.; Hoffmann, R.J.; Moore, J.D.; Brown, W.M.
submitted to the EMBL Data Library, June 1992
A:Reference number: S66600
A:Accession: S66600
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-311 <HOF>
A:Cross-references: EMBL:M33760; NID:G342485; PIDN:AAA31307.1; PID:G342488
R.; Hoffmann, R.J.; Moore, J.D.; Brown, W.M.
Genetics 131, 397-412, 1992
A:Title: A novel mitochondrial genome organization for the blue mussel, Mytilus edulis.
A:Reference number: S28743; MUID:92354892; PMID:1386586
A:Accession: S28753
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-163, N', 165-311 <BRO>
A:Cross-references: EMBL:M33760
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGCA
C:Superfamily: cytochrome-c oxidase chain III
C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane

C/Accession: H8532
 R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A/Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
 A/Reference number: A85480; MUID:21074935; PMID:11206551
 A/Accession: H8532
 A/Status: Preliminary
 A/Molecule type: DNA
 A/Residues: 1-364 <STO>
 A/Cross-references: GB:AE005174; NID:912513218; PIDN:AAG54724.1; GSPDB:GN00145; UWGP:Z04
 A/Experimental source: strain O157:H7, substrain BDL933
 C/Genetics:
 A/Gene: yaiW

Query Match 55.3%; Score 42; DB 2; Length 364;
 Best Local Similarity 70.0%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 DINPAWYASR 12
 | | | | |
 Db 233 DFNAGWYASR 242

RESULT 13

AR0548
 Probable lipoprotein yaiW [imported] - *Salmonella enterica* subsp. *enterica* serovar Typhi
 C/Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
 A/Note: this species has also been called *Salmonella typhi*
 C/Accession: AH0548
 A/Status: Preliminary
 A/Molecule type: DNA
 A/Residues: 1-364 <PAR>
 A/Cross-references: GB:AL513382; PIDN:CAD08832.1; PID:G16501646; GSPDB:GN00176
 C/Genetics:
 A/Gene: yaiW

Query Match 55.3%; Score 42; DB 2; Length 364;
 Best Local Similarity 70.0%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 DINPAWYASR 12
 | | | | |
 Db 233 DFNAGWYASR 242

RESULT 14

G83309
 conserved hypothetical protein PA2695 [imported] - *Pseudomonas aeruginosa* (strain PAO1)
 C/Species: *Pseudomonas aeruginosa*
 C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C/Accession: G83309
 R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A/Title: Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic patho
 A/Reference number: AB2950; MUID:20437337; PMID:10984043
 A/Accession: G83309
 A/Status: Preliminary
 A/Molecule type: DNA
 A/Residues: 1-367 <STO>
 A/Cross-references: GB:AE004697; GB:AE004091; NID:99948760; PIDN:AAGC6083.1; GSPDB:GN001
 A/Experimental source: strain PAO1

C/Genetics:
 A/Gene: PA2695

Query Match 55.3%; Score 42; DB 2; Length 367;
 Best Local Similarity 70.0%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 DINPAWYASR 12
 | | | | |
 Db 232 DFNAGWYASR 241

RESULT 15

G64720
 Probable amino acid transport protein yaaJ, sodium-dependent - *Escherichia coli* (strain F
 C/Species: *Escherichia coli*
 C/Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
 C/Accession: G64720
 R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A/Title: The complete genome sequence of *Escherichia coli* K-12.
 A/Reference number: A64720; MUID:97426617; PMID:9278503
 A/Accession: G64720
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-476 <BLAT>
 A/Cross-references: GB:AE000111; GB:U00096; NID:G1786181; PIDN:AAC73118.1; PID:G1786188;
 A/Experimental source: strain K-12, substrain M61655
 C/Genetics:
 A/Gene: yaaJ

Query Match 55.3%; Score 42; DB 2; Length 476;
 Best Local Similarity 47.1%; Pred. No. 30;
 Matches 8; Conservative 2; Mismatches 1; Indels 6; Gaps 1;

QY 3 DIN-----PAWYASRG 13
 | | | | |
 Db 120 DVNGQFRGPGAWYARG 136

Search completed: August 12, 2004, 14:50:27
 Job time : 6.74419 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 3.40116 Seconds
(without alignments)
199.024 Million cell updates/sec

Title: US-09-700-643A-2_COPY_12_24

Perfect score: 76

Sequence: 1 TPDINPAWASRG 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	87	1 PRRP_HUMAN	P81277 homo sapien
2	72	94.7	98	1 PRRP_BOVIN	P81264 bos taurus
3	68	89.5	83	1 PRRP_RAT	P81278 rattus norv
4	42	55.3	364	1 YAAJ_ECOLI	P77562 escherichia
5	42	55.3	476	1 YAAJ_ECOLI	P30143 escherichia
6	42	55.3	512	1 UB33_SCHPO	O94269 schizosacch
7	41	53.9	986	1 CYGR_ARBPU	P11528 arabacia pun
8	40	52.6	208	1 YK65_CABEL	P34347 caenorhabdi
9	40	52.6	328	1 CNA4_HUMAN	Q9nvrs homo sapien
10	40	52.6	376	1 OPS1_LIMPO	P35360 limulus pol
11	40	52.6	376	1 OPS2_LIMPO	P35361 limulus pol
12	40	52.6	626	1 THIC_PSEPK	Q88da5 pseudomonas
13	40	52.6	757	1 IMMT_MOUSE	Q8cag8 mus musculu
14	39	51.3	678	1 AMM1_MOUSE	Q9jnd6 rhodobacter
15	38.5	50.7	240	1 PLSC_HELPY	Q25903 helicobacte
16	38	50.0	197	1 I18B_HUMAN	Q95998 homo sapien
17	38	50.0	360	1 WNT2_HUMAN	P09544 homo sapien
18	38	50.0	378	1 OPSD_CAMAB	Q17292 camponotus
19	38	50.0	378	1 OPSD_CATBO	Q17296 cataglyphis
20	38	50.0	385	1 YE1B_ECOLI	P25747 escherichia
21	38	50.0	402	1 EX7L_STRCO	Q9fbm3 streptomyce
22	38	50.0	455	1 Y755_MYCTU	P71555 mycobacteri
23	38	50.0	467	1 RFAL_CRIFA	Q23696 crithidia f
24	38	50.0	503	1 ANXB_BOVIN	P27214 bos taurus
25	38	50.0	3083	1 POLG_ZYMYR	Q89330 z genome po
26	37.5	49.3	772	1 LMBT_HUMAN	Q9y468 homo sapien
27	37	48.7	61	1 VLCE_ECOLI	P77087 escherichia
28	37	48.7	91	1 RL31_SYNEL	Q8dmk6 synchococc
29	37	48.7	212	1 PCP_STAAR	Q53596 staphylococ
30	37	48.7	284	1 AMPE_ECOLI	P13017 escherichia
31	37	48.7	290	1 YEAB_BACSU	P46348 bacillus su
32	37	48.7	295	1 TRPI_PSEAE	P11720 pseudomonas
33	37	48.7	310	1 TNFC_MARMO	Q9jml0 marmota mon

34	37	48.7	315	1 OAA4_HUMAN	Q9h209 homo sapien
35	37	48.7	369	1 Y264_SYNY3	P73879 synchocyst
36	37	48.7	393	1 CYGR_CHRVI	O82947 chromatium
37	37	48.7	454	1 YQSI_CABEL	Q09109 caenorhabdi
38	37	48.7	511	1 CP12_CANFA	P56592 canis famli
39	37	48.7	581	1 POL_MLVK	P31795 radiatation m
40	37	48.7	605	1 PALZ_ECOLI	P21517 escherichia
41	37	48.7	697	1 LCPF_HUMAN	Q9uku0 homo sapien
42	37	48.7	812	1 PDA1_ORVSA	Q43007 oryza sativ
43	37	48.7	843	1 POL_MLVAK	P03357 akr murine
44	37	48.7	885	1 CHE_SERMA	O54468 serratia ma
45	37	48.7	1036	1 YAN2_SCHPO	Q10068 schizosacch

ALIGNMENTS

RESULT 1

PRRP_HUMAN	STANDARD;	PRT;	87 AA.
AC	P81277;		
DT	30-MAY-2000 (Rel. 39, Created)		
DT	30-MAY-2000 (Rel. 39, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-releasing peptide PrRP20].		
DE	PRH.		
GN	Homo sapiens (Human).		
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	SEQUENCE FROM N.A.		
RP	TISSUE=Brain;		
RC	MEDLINE=982668781; PubMed=9607765;		
RA	Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,		
RA	Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,		
RA	Kurokawa T., Nishimura O., Onda H., Fujino M.,		
RA	"A prolactin-releasing peptide in the brain."		
RL	Nature 393:272-276(1998).		
RN	[2]		
RP	TISSUE SPECIFICITY.		
RX	MEDLINE=99428652; PubMed=10498338;		
RA	Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,		
RA	Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,		
RA	Sumino Y., Fujino M.,		
RT	"Tissue distribution of prolactin-releasing peptide (PrRP) and its receptor";		
RT	Regul. Pept. 83:1-10(1999)		
RL	FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.		
CC	TISSUE SPECIFICITY: Medulla oblongata and hypothalamus.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
CC	EMBL; AB015419; BAA29027.1; --		
DR	MIM; 602663;		
DR	GO; GO:0005180; F:peptide hormone; TAS.		
DR	Hormone; Amidation; Signal.		
FT	SIGNAL 1 22		
FT	PEPTIDE 23 53		
FT	PEPTIDE 34 53		
FT	PROPEP 56 87		
FT	MOD_RES 53 53		
SQ	SEQUENCE 87 AA; 9639 MW; 229A2F3F50CF981B CRC64;		

DE Probable ubiquitin carboxyl-terminal hydrolase 3 (EC 3.1.2.15)
DE (Ubiquitin thiolesterase 3) (Ubiquitin-specific processing protease 3)
DE (Deubiquitinating enzyme 3).
GN UBP3 OR SPBP87.21.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sourd V., Williams J., Peat N., Hayes J., Baker S., Baeham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Muncall K., Murphy L., Niblett D., Odell J.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkstraet G., Aert R., Robben J., Grymonprez B.,
RA Weltjens J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Pritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambolt R., Purnelle B.,
RA Coffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -!- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2O) =
CC ubiquitin + a thiol.
CC -!- SIMILARITY: Belongs to peptidase family C19.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AL032684; CAA21806.1; -.
DR PIR: T40815; T40815.
DR GeneDB Spombe; SPBP87.21; -.
DR GO: GO:0006464; P:protein modification; ISS.
DR GO: GO:0006508; P:proteolysis and peptidolysis; ISS.
DR InterPro: IPR001394; Peptidase_C19.
DR Pfam: Pf00443; UCH; 1.
DR PROSITE: PS00972; UCH_2_1; 1.
DR PROSITE: PS00973; UCH_2_2; 1.
DR PROSITE: PS02035; UCH_2_3; 1.
KW Ub1 conjugation pathway; Hydrolase; Thiol protease; Multigene family.
FT ACT_SITE 142 142 BY SIMILARITY.
FT ACT_SITE 445 445 BY SIMILARITY.
FT ACT_SITE 453 453 BY SIMILARITY.
SQ SEQUENCE 512 AA; 58081 MW; 1F97DA2C7720695D CRC64;

Query Match 55.3%; Score 42; DB 1; Length 512;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 2 PDINPANYASRG 13
|:|:|:|:|:|

DB 361 PEILPEWHSSKG 372
RESULT 7
CYGR ARBPU
ID CYGR ARBPU STANDARD; PRT; 986 AA.
AC P11528;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Resact receptor precursor (Guanylate cyclase) (EC 4.6.1.2).
OS Arbacia punctulata (Punctuate sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoida; Euechinozoa; Echinidea; Arbacoidea; Arbaciidae; Arbacia.
OX NCBI_TaxID=7641;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Testis;
RC MEDLINE=98318927; PubMed=2901039;
RX Singh S., Lowe D.G., Thorpe D.S., Rodriguez H., Kuang W.-J.,
RA Dangott L.J., Chinkers M., Goeddel D.V., Garbers D.L.;
RT "Membrane guanylate cyclase is a cell-surface receptor with homology
RT to protein kinases."
RL Nature 334:708-712(1988).
CC -!- FUNCTION: IMPLICATED AS A CELL-SURFACE RECEPTOR ON SPERMATOZOEA
CC FOR 'RESACT' A CHEMOTACTIC PEPTIDE, AND ON VARIOUS OTHER CELLS
CC AS A RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE.
CC -!- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Contains 1 protein kinase-like domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X12874; CAA31367.1; -.
DR PIR: S05480; OXUGA.
DR InterPro: IPR001828; ANF_receptor.
DR InterPro: IPR001054; G_cyclase.
DR Pfam: PF01094; ANF_receptor; 1.
DR Pfam: PF00211; guanylate_cyc; 1.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
KW Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;
KW cAMP biosynthesis; Signal.
FT CHAIN 1 21
FT DOMAIN 22 986 RESACT RECEPTOR
FT DOMAIN 22 507 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 529 528 POTENTIAL.
FT DOMAIN 529 986 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 568 836 PROTEIN KINASE LIKE.
FT CARBOHYD 185 185 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 361 361 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 410 410 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 986 AA; 111284 MW; B40238A74CCAF52 CRC64;

Query Match 53.9%; Score 41; DB 1; Length 986;
Best Local Similarity 62.5%; Pred. No. 38;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 PDINPAWY 9
|:|:|:|:|:|

DB 475 PDINPVDH 482

RESULT 8
YK69_CAEEL

ID YK69 CAEEL STANDARD; PRT; 208 AA.
AC P34347;
DT 01-FEB-1994 (Rel. 28, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
GN Hypothetical protein C29E4.9 in chromosome III.
GN C29E4.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Chordata; Rhabdita; Rhabditoidea;
OC Rhabditiidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsley T., Cooper A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kersey J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RN REVISIONS.
RA Waterston R.;
RA Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; L23651; AAA27956.2; -;
DR PIR; S44765; S44765.
DR WormPep; C29E4.9; CE29199.
DR Hypothetical protein.
SQ SEQUENCE 208 AA; 23729 MW; FD4A57F6609EEB32 CRC64;
Query Match 52.6%; Score 40; DB 1; Length 208;
Best Local Similarity 61.5%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 TPDPINPAWASRG 13
Db 91 TEVNPAPAAQRG 103
RESULT 9
CNA4 HUMAN STANDARD; PRT; 328 AA.
ID CNA4_HUMAN
AC Q9NVR5; Q86TV8; Q96925;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein C14orf104.
GN C14ORF104.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RP Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,

RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA Wakamatsu A., Nakamura Y., Nagahari K., Masuhio Y., Oshima A.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Neuroblastoma;
RA Li W.B., Gruber C., Jesse J., Polayes D.;
RT "Full-length cDNA libraries and normalization.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RX PubMed=12508121;
RA Heilig R., Eckenberg R., Petit J.-L., Fonknechten N., Da Silva C.,
RA Cattolico L., Levy M., Barbe V., De Bernardinis V., Ureta-Vidal A.,
RA Pelletier E., Vico V., Anthonard V., Rowen L., Madan A., Qin S.,
RA Sun H., Du H., Pepin K., Artiguenave F., Robert C., Cruaud C.,
RA Bruls T., Jallion O., Friedlander L., Samson G., Brottier P.,
RA Cure S., Segrens B., Aniere F., Samain S., Crespeau H., Abbasi N.,
RA Alich N., Boscus D., Dickhoff R., Dors M., Dubois I., Friedmann S.,
RA Gouyvenoux M., James R., Madan A., Maitey-Estrada B., Mangenot S.,
RA Martins N., Menard M., Ocas S., Ratcliffe A., Shaffer T., Trask B.,
RA Vacherie B., Bellemere C., Belser C., Beshard-Gonnet M.,
RA Bartol-Mavel D., Boutard M., Briez-Silla S., Combette S.,
RA Dufosse-Laurent V., Ferron C., Lechaplais C., Lousse S., Muselet D.,
RA Magdelenat G., Pateau E., Petit E., Sirivan-Trukiewicz P., Trybou A.,
RA Vega-Czarny N., Bataille E., Bluet E., Bordelais I., Dubois M.,
RA Dumont C., Guerin T., Haffray S., Hammadi R., Muanga J., Pellouin V.,
RA Robert D., Wunderle E., Gauguet G., Roy A., Sainte-Marthe L.,
RA Verdier J., Verdier-Discala C., Hillier L.W., Fulton L., McPherson J.,
RA Matsuda F., Wilson R., Scarpelli C., Gyapay G., Wincker P., Saurin W.,
RA Quetier F., Waterston R., Hood L., Weissenbach J.;
RT "The DNA sequence and analysis of human chromosome 14.";
RL Nature 421:601-607(2003).
RN [4]
RN SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Bone marrow, and Colon;
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9NVR5-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9NVR5-2; Sequence=VSP_008390;
CC Note=May be due to exon skipping;
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL; AK001425; BA91684.1; -;
 DR EMBL; BX248264; CAD62592.1; ALT INIT.
 DR EMBL; AL139099; -; NOT_ANNOTATED_CDS.
 DR EMBL; BC013322; AAH13322.1; -;
 DR EMBL; BC011400; AAH11400.1; -;
 DR Genew; HGNC:20188; C14orf104.
 DR Alternative splicing.
 KW VARSPIC 113 160 Missing (in isoform 2).
 FT /FtId=VSP_008390.
 FT
 SQ SEQUENCE 328 AA; 36943 MW; 87875CE2A89AF663 CRC64;
 Query Match 52.6%; Score 40; DB 1; Length 328;
 Best Local Similarity 60.0%; Pred. No. 18;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 3 DINPAWYASR 12
 | : | | | | |
 Db 42 DLNPLMYKLR 51

 RESULT 10
 OPS1 LIMPO STANDARD; PRT; 376 AA.
 AC P35360;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Lateral eye opsin.
 OS Limulus polyphemus (Atlantic horseshoe crab).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;
 OC Limulidae; Limulus.
 OX NCBI_TaxID=6850;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lateral eye;
 RX MEDLINE=93317641; PubMed=8327495;
 RA Smith W.C., Price D.A., Greenberg R.M., Battelle B.-A.;
 RT "Opsins from the lateral eyes and ocelli of the horseshoe crab,
 Limulus polyphemus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:6150-6154 (1993).
 CC -!- FUNCTION: Visual pigments are the light-absorbing molecules that
 mediate vision. They consist of an apoprotein, opsin, covalently
 linked to cis-retinal.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: LATERAL EYE.
 CC -!- PTM: Some or all of the carboxyl-terminal Ser or Thr residues may
 be phosphorylated (by similarity).
 CC -!- MISCELLANEOUS: THIS OPSIN HAS AN ABSORPTION MAXIMUM AT 520 NM.
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC Opsin subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL; L03791; AAA28273.1; -;
 DR EMBL; L03781; AAA02498.1; -;
 DR PIR; B48197; B48197.
 DR HSSP; P02699; 1EDV.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR InterPro; IPR001760; Opsin.
 DR Pfam; PF00001; 7tm1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
 DR PROSITE; PS02062; G-PROTEIN_RECEP_F1_2; 1.
 DR PROSITE; PS00238; OPSIN; 1.

KW Photoreceptor; Retinal protein; Transmembrane; Glycoprotein; Vision;
 KW Phosphorylation; G-protein coupled receptor.
 FT DOMAIN 1 46
 FT TRANSMEM 47 71
 FT DOMAIN 72 83
 FT TRANSMEM 84 108
 FT DOMAIN 109 123
 FT TRANSMEM 124 143
 FT DOMAIN 144 162
 FT TRANSMEM 163 186
 FT DOMAIN 187 210
 FT TRANSMEM 211 238
 FT DOMAIN 239 274
 FT TRANSMEM 275 298
 FT DOMAIN 299 306
 FT TRANSMEM 307 331
 FT DOMAIN 332 376
 FT DISULFID 120 197
 FT BINDING 318 318
 FT CARBOHYD 17 17
 FT CARBOHYD 193 193
 SQ SEQUENCE 376 AA; 42139 MW; CCE401766AB06F26 CRC64;
 Query Match 52.6%; Score 40; DB 1; Length 376;
 Best Local Similarity 50.0%; Pred. No. 21;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 2 PDINPAWYASRG 13
 | : | | | | |
 Db 40 PPMNPLMYSLG 51

 RESULT 11
 OPS2 LIMPO STANDARD; PRT; 376 AA.
 AC P35361;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ocellar opsin.
 OS Limulus polyphemus (Atlantic horseshoe crab).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;
 OC Limulidae; Limulus.
 OX NCBI_TaxID=6850;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Median ocelli;
 RX MEDLINE=93317641; PubMed=8327495;
 RA Smith W.C., Price D.A., Greenberg R.M., Battelle B.-A.;
 RT "Opsins from the lateral eyes and ocelli of the horseshoe crab,
 Limulus polyphemus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:6150-6154 (1993).
 CC -!- FUNCTION: Visual pigments are the light-absorbing molecules that
 mediate vision. They consist of an apoprotein, opsin, covalently
 linked to cis-retinal.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: OCELLAR CELLS; MEDIAN OCELLI.
 CC -!- PTM: Some or all of the carboxyl-terminal Ser or Thr residues may
 be phosphorylated (by similarity).
 CC -!- MISCELLANEOUS: THIS OPSIN HAS AN ABSORPTION MAXIMUM AT 530 NM.
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC Opsin subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL; L03792; AAA28274.1; -;
 DR EMBL; L03782; AAA02499.1; -;

DR TIGR; HPI348; -
DR InterPro; IPR002123; Acyltransferase.
DR InterPro; IPR004552; AGP acyltrn.
DR Pfam; PF01553; Acyltransferase; 1.
DR SMART; SM00563; PlsC; 1.
DR TIGRFAMs; TIGR00530; AGP acyltrn; 1.
KW Phospholipid biosynthesis; Transferase; Acyltransferase;
KW Inner membrane; Complete proteome.
SQ SEQUENCE 240 AA; 27745 MW; 22BD5D0EB190BDD CRC64;

Query Match 50.7%; Score 38.5; DB 1; Length 240;
Best Local Similarity 70.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 TPDIN-PAWY 9
Db 208 TPDNSPTWY 217

Search completed: August 12, 2004, 14:44:52
Job time : 3.4016 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 17,9128 Seconds
(without alignments)
228.984 Million cell updates/sec

Title: US-09-700-643A-2_COPY_12_24
Perfect score: 76
Sequence: 1 TPDINPAWYASRG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rviro:*

16: sp_bacteriap:*

17: sp_archaeap:*

17	43	56.6	402	2	Q9KIW3
18	43	56.6	614	10	Q8GZ21
19	43	56.6	1043	16	Q8A693
20	43	56.6	1093	3	Q8XOR0
21	43	56.6	1359	10	Q8SW46
22	43	56.6	1361	10	Q8GUQ8
23	43	56.6	1364	10	Q8SW45
24	42.5	55.9	664	16	Q911W4
25	42.5	55.9	779	3	Q9P5J9
26	42	55.3	72	8	Q9TD69
27	42	55.3	120	8	Q8SR31
28	42	55.3	221	16	Q8RTS3
29	42	55.3	338	5	Q9B5L8
30	42	55.3	350	16	Q8GB99
31	42	55.3	358	16	Q882Y3
32	42	55.3	363	16	Q8PPU3
33	42	55.3	363	16	Q8P501
34	42	55.3	364	16	Q8ZRF4
35	42	55.3	364	16	Q8X5H6
36	42	55.3	364	16	Q8FKES
37	42	55.3	364	16	Q8Z8Z7
38	42	55.3	364	16	Q83MA5
39	42	55.3	367	16	Q910E7
40	42	55.3	419	13	Q7ZTL7
41	42	55.3	467	11	Q8R054
42	42	55.3	468	11	Q8XLF8
43	42	55.3	476	16	Q8XA75
44	42	55.3	501	16	Q88B54
45	42	55.3	505	16	Q88RQ2

ALIGNMENTS

RESULT 1

Q8WN12

ID Q8WN12 PRELIMINARY; PRT; 98 AA.

AC Q8WN12;

DT 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Preprolactin-releasing peptide.

OS Ovis aries (Sheep).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Caprinae; Ovis.

OX NCBI_TaxID=9940;

RN [1]

RP SEQUENCE FROM N.A.

RA Curlewis J.D., Kusters D.H.L., Barclay J.L., Anderson S.T.;

RT "Prolactin-releasing peptide (prp) in the ewe: cDNA cloning, mRNA

RT distribution and effects on prolactin secretion in vitro and in

RT vivo.";

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF450453; AAL47178.1; -.

SQ SEQUENCE 98 AA; 10513 MW; 2A53331ED62CAAB5 CRC64;

Query Match	94.7%	Score 72;	DB 6;	Length 98;
Best Local Similarity	92.3%	Pred. No. 0.00017;		
Matches 12;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	TPDINPAWYASRG 13		
Db	34	TPDINPAWYAGRG 46		
RESULT 2				
Q9W624				
ID Q9W624		PRELIMINARY; PRT; 117 AA.		
AC Q9W624;				
DT 01-NOV-1999 (TREMBlrel. 12, Created)				
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)				
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)				

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Match Length DB ID. Description

1 72 94.7 98 6 Q8WN12 Q8wn12 ovis aries

2 49 64.5 117 13 Q9W624 Q9w624 carassius a

3 47 61.8 54 4 Q7Z6Y1 Q7z6y1 homo sapien

4 47 61.8 465 4 Q60687 Q60687 homo sapien

5 47 61.8 465 4 Q8W85 Q8w85 homo sapien

6 45 59.2 179 5 Q20170 Q20170 caenorhabdi

7 45 59.2 816 10 Q7XMH8 Q7xmh8 oryza sativ

8 44 57.9 387 16 Q7VUL5 Q7vul5 bordetella

9 44 57.9 432 16 Q7WQX8 Q7wgx8 bordetella

10 44 57.9 432 16 Q7W200 Q7w200 bordetella

11 44 57.9 527 5 Q76383 Q76383 caenorhabdi

12 44 57.9 813 17 Q8ZXH2 Q8zxh2 pyrobaculum

13 43 56.6 186 16 Q7U4H3 Q7u4h3 synechococc

14 43 56.6 140 2 Q93MG0 Q93mg0 thioabacillu

15 43 56.6 177 2 Q93MG5 Q93mg5 thioabacillu

16 43 56.6 341 16 Q7VM68 Q7vm68 haemophilus

Q9kiw3 thioabacillu

Q8g221 arabisdopsi

Q8a693 bacteroides

Q8x0r0 neurospora

Q9sw46 arabisdopsi

Q8puq8 arabisdopsi

Q9sw45 arabisdopsi

Q9p5j9 neurospora

Q9td69 mytilus cal

Q8sr31 mytilus tro

Q9rts3 deinococcus

Q9bl18 trichinella

Q8g9b9 pseudomonas

Q882y3 pseudomonas

Q88pu3 xanthomonas

Q8p501 xanthomonas

Q8rf4 salmoneella

Q8x5h6 escherichia

Q8fke5 escherichia

Q8a27 salmoneella

Q8ma5 shigella fl

Q910e7 pseudomonas

Q7zt17 xenopus lae

Q8r054 mus musculu

Q8k1f8 mus musculu

Q8a75 escherichia

Q8b54 pseudomonas

Q88rq2 pseudomonas

DE C-RF amide.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Satake H., Kinakata H., Fujimoto M.;
RT "Carassius Rfanide (C-RF amide).";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB020024; BAA76662.1; -.
SQ SEQUENCE 117 AA; 12879 MW; D5DC4CB22038C2B0 CRC64;
Query Match 64.5%; Score 49; DB 13; Length 117;
Best Local Similarity 53.8%; Pred. No. 1.5;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Qy 1 TPDINPAWYASRG 13
Db 56 SPEIDPFYVGRG 68
RESULT 3
Q7Z6Y1 ID Q7Z6Y1 PRELIMINARY; PRT; 54 AA.
AC Q7Z6Y1;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE DJ479J7.3 (Sushi-repeat protein (SRPUL)) (Fragment).
GN DJ479J7.3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lawlor S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL035608; CAB55682.1; -.
FT NON TER 54
SQ SEQUENCE 54 AA; 6110 MW; E2F3C39F7B961A9F CRC64;
Query Match 61.8%; Score 47; DB 4; Length 54;
Best Local Similarity 53.8%; Pred. No. 1.4;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 1 TPDINPAWYASRG 13
Db 18 TPAVPTWYAGSG 30
RESULT 4
O60687 ID O60687 PRELIMINARY; PRT; 465 AA.
AC O60687;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Sushi-repeat protein (Sushi-repeat containing protein).
GN SRPUL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kurosawa H., Inukai T., Inaba T., Inaba K.-S., Sinyo T.,
RA Rakesraw K.M., Naeve C.W., Look T.A.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]

RP SEQUENCE FROM N.A.
RA Huang C.-H., Chen H., Peng J., Chen Y.;
RT "Cloning and characterization of the sushi-repeat containing protein
RT (SRP) as a novel interaction partner of Rh type C glycoprotein
RT (RnCG).";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF060567; AAC15765.1; -.
DR EMBL; AF93649; AAW73693.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR003410; Hyalin.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF02494; HVR; 1.
DR Pfam; PF00084; sushi; 3.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
SQ SEQUENCE 465 AA; 52971 MW; 4D752B187FF3EFB8 CRC64;
Query Match 61.8%; Score 47; DB 4; Length 465;
Best Local Similarity 53.8%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 1 TPDINPAWYASRG 13
Db 18 TPAVPTWYAGSG 30
RESULT 5
Q8W85 ID Q8W85 PRELIMINARY; PRT; 465 AA.
AC Q8W85;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Sushi-repeat protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020733; AAH20733.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR003410; Hyalin.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF02494; HVR; 1.
DR Pfam; PF00084; sushi; 3.
DR SMART; SM00032; CCP; 3.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
SQ SEQUENCE 465 AA; 52957 MW; 3D7229487DA1B8BD CRC64;
Query Match 61.8%; Score 47; DB 4; Length 465;
Best Local Similarity 53.8%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 1 TPDINPAWYASRG 13
Db 18 TPAVPTWYAGSG 30
RESULT 6
Q20170 ID Q20170 PRELIMINARY; PRT; 179 AA.
AC Q20170;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE C.elegans WRT-3 protein (Corresponding sequence F38E11.7).
GN WRT-3.
OS Caenorhabditis elegans.

```

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RA "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Matthews P.;
RA Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RL EMBL; Z68342; CAA92775.2; -.
SQ SEQUENCE 179 AA; 20965 MW; F9FA6891A2BF2BDD CRC64;

Query Match 59.2%; Score 45; DB 5; Length 179;
Best Local Similarity 75.0%; Pred. No. 11;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PDINPAWY 9
Db 159 PDVKNPAWY 166

RESULT 7
ID Q7XMH8 PRELIMINARY; PRT; 816 AA.
AC Q7XMH8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CSJNB0028123.4 protein.
DE OSJNB0028123.4.
GN CSJNB0028123.4.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Fu G., Wang S.Y., Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F.,
RA Jia J., Yin H.F., Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y.,
RA Shao Y., Sun Y., Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W.,
RA Sheng H.H., Gu J.L., Chen S.T., Ni L., Zhu F.H., Han B., Feng Q.,
RA Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X., Liu Y.L., Mu J., Yu Z.,
RA Chen L., Fan D.L., Weng Q.J., Zhang L., Lu Y.Q., Yu S.L., Liu X.H.,
RA Lu T.T., Zhang Y.J., Lu Y., Li C., Li T., Zhang Y., Hu H., Jia P.X.,
RA Qian Y.M., Ying K., Zhou B., Chen Z.H., Hao P., Zhang L., Wu M.,
RA Zhang R.Q., Guan J.P., Hong G.F.
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL662944; CA504622.1; -.
SQ SEQUENCE 816 AA; 91136 MW; 7ED04CAF2E989700 CRC64;

Query Match 59.2%; Score 45; DB 10; Length 816;
Best Local Similarity 63.6%; Pred. No. 56;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 PDINPAWYAGR 12
Db 72 PKLNPAWYAGR 82

RESULT 8
ID Q7VUL5 PRELIMINARY; PRT; 387 AA.
AC Q7VUL5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

```

```

DE S-adenosylmethionine synthetase (EC 2.5.1.16).
GN METK OR Bp3071.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama 1 / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jageis K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640420; CAE43340.1; -.
SQ SEQUENCE 387 AA; 42002 MW; 91F207251B1A4C94 CRC64;

Query Match 57.9%; Score 44; DB 16; Length 387;
Best Local Similarity 64.3%; Pred. No. 37;
Matches 9; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 1 TPDINPA--WYASR 12
Db 135 TFDLMPAPIWYAGR 148

RESULT 9
Q7WQX8
ID Q7WQX8 PRELIMINARY; PRT; 432 AA.
AC Q7WQX8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE S-adenosylmethionine synthetase (EC 2.5.1.16).
GN METK OR B80195.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jageis K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640437; CAE30694.1; -.
SQ SEQUENCE 432 AA; 46834 MW; D86B832A85FCA9D CRC64;

Query Match 57.9%; Score 44; DB 16; Length 432;
Best Local Similarity 64.3%; Pred. No. 42;
Matches 9; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

```

```

Qy 1 TPDINPA--WYASR 12
    |||: || |||||
Db 180 TPDLMPIWYAH 193

RESULT 10
Q7W200 PRELIMINARY; PRT; 432 AA.
AC Q7W200;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE S-adenosylmethionine synthetase (EC 2.5.1.6).
GN METK OR BPP0192.
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP "SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cardeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norbertczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640423; CAE39933.1; -.
KW Transferrase; Complete proteome.
SQ SEQUENCE 432 AA; 46834 MW; D86B8323A95FCA9D CRC64;

Query Match 57.9%; Score 44; DB 16; Length 432;
Best Local Similarity 64.3%; Pred. No. 42;
Matches 9; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Qy 1 TPDINPA--WYASR 12
    |||: || |||||
Db 180 TPDLMPIWYAH 193

RESULT 11
Q76383 PRELIMINARY; PRT; 527 AA.
AC Q76383;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN C24G6.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP "SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP "SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;

```

```

RA Greco T., Bradshaw H., Keppler D.;
RT "The sequence of C. elegans cosmid C24G6.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF067936; AAC19213.1; -.
DR F01; T33175; T33175.
DR WormPep; C24G6.6; CE17462.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR002937; Amino oxidase.
DR InterPro; IPR002005; NAD B5.
DR Pfam; PF01593; Amino oxidase; 1.
KW Hypothetical protein.
SQ SEQUENCE 527 AA; 59805 MW; 9FBB1FB84437C5CB CRC64;

Query Match 57.9%; Score 44; DB 5; Length 527;
Best Local Similarity 58.3%; Pred. No. 52;
Matches 7; Conservative 2; Mismatches 3; Indels 3; Gaps 0;

Qy 2 PDINPAWYASRG 13
    |||: |||||
Db 370 PNLSAWYASRG 381

RESULT 12
Q8ZXH2 PRELIMINARY; PRT; 813 AA.
AC Q8ZXH2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Malate synthase (glcB).
GN PAE1287.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792859;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009813; AAL63376.1; -.
DR GO; GO:0004474; F:malate synthase activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006097; P:glyoxylate cycle; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR GO; GO:0006099; P:tricarboxylic acid cycle; IEA.
DR InterPro; IPR000566; Lipoclin_cytfabp.
DR InterPro; IPR001485; Malate synthase.
DR Pfam; PF01274; Malate synthase; 1.
DR PROSITE; PS00213; LIPOCALIN; 1.
KW Complete proteome.
SQ SEQUENCE 813 AA; 93836 MW; 1EB05EA41EA06FE8 CRC64;

Query Match 57.9%; Score 44; DB 17; Length 813;
Best Local Similarity 63.6%; Pred. No. 83;
Matches 7; Conservative 1; Mismatches 3; Indels 3; Gaps 0;

Qy 3 DINPAWYASRG 13
    |||: |||||
Db 152 DASPAWYIPRG 162

RESULT 13

```

Q7U4H3
ID Q7U4H3 PRELIMINARY; PRT; 86 AA.
AC Q7U4H3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 50S ribosomal protein L31.
GN RPL31, RPL31, RPL31 OR SYN2095.
OS Synechococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
CX NCBI_TaxID=84588;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825697; PubMed=12917641;
RA Palenik B., Brahamsha B., Laxner F.W., Land M., Hauser L., Chain P.,
RA Lander J., Regala W., Allen E.B., McCarren J., Paulsen I.,
RA Dufresne A., Patensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synechococcus.";
RL Nature 424:1037-1042(2003).
DR EMBL; BX569694; CAB08610.1;
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 86 AA; 9537 MW; 9B88EA9CC2EE34B7 CRC64;
Query Match 56.6%; Score 43; DB 16; Length 86;
Best Local Similarity 75.0%; Pred. No. 11;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 2 PDINPANY 9
Db 4 PDHPTWY 11
RESULT 14
Q93MGO PRELIMINARY; PRT; 140 AA.
ID Q93MGO;
AC Q93MGO;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative cytochrome b (Fragment).
GN CYTB OR PETB1.
OS Thiobacillus ferrooxidans.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Acidithiobacillales;
OC Acidithiobacillaceae; Acidithiobacillus.
CX NCBI_TaxID=920;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC33020;
RA Levine G., Bruscella P., Guacurano M., Inostroza C., Jedlicki E.,
RA Bonnefoy V., Holmes D.S.;
RT "Characterization of the pet and res operons of Acidithiobacillus
ferrooxidans.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AJ318506; CAC4964.1;
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; cytochrome_b_c_1.
KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 140
FT NON_TER 140 140

SQ SEQUENCE 140 AA; 15881 MW; 01DB8822ACE9F1BC CRC64;
Query Match 56.6%; Score 43; DB 2; Length 140;
Best Local Similarity 60.0%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 2 PDINPANYAS 11
Db 85 PDVTPPWYLS 94
RESULT 15
Q93MGS PRELIMINARY; PRT; 177 AA.
ID Q93MGS;
AC Q93MGS;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative cytochrome b (Fragment).
GN CYTB OR PETB2.
OS Thiobacillus ferrooxidans.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Acidithiobacillales;
OC Acidithiobacillaceae; Acidithiobacillus.
CX NCBI_TaxID=920;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC33020;
RA Bruscella P., Levine G., Ratouchniak J., Holmes D., Bonnefoy V.;
RT "A second operon encoding a bcl complex in Acidithiobacillus
ferrooxidans.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AJ318501; CAC4960.1;
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; cytochrome_b_c_1.
KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 177
FT NON_TER 177 177
SQ SEQUENCE 177 AA; 20264 MW; D841E0DD8F878882 CRC64;
Query Match 56.6%; Score 43; DB 2; Length 177;
Best Local Similarity 60.0%; Pred. No. 24;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 2 PDINPANYAS 11
Db 88 PDVTPPWYLS 97
Search completed: August 12, 2004, 14:49:02
Job time : 18.9128 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 27.5872 Seconds
(without alignments)
133.146 Million cell updates/sec

Title: US-09-700-643A-2_COPY_12_24
Perfect score: 76
Sequence: 1 TPDINPAWASRG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	14	AA49300	19P2 liga
2	76	100.0	20	AAW31394	Human typ
3	76	100.0	20	AAW97236	Human typ
4	76	100.0	20	AAW49294	19P2 liga
5	76	100.0	20	AAW10365	Human oxy
6	76	100.0	20	AAW90992	Prolactin
7	76	100.0	20	AAW62534	Human CRH
8	76	100.0	20	AAW26404	Human PrR
9	76	100.0	20	AAW60846	Peptide p
10	76	100.0	21	AAW31395	Human typ
11	76	100.0	21	AAW10366	Human oxy
12	76	100.0	21	AAW62535	Human CRH
13	76	100.0	21	AAW60847	Peptide p
14	76	100.0	22	AAW31396	Human typ
15	76	100.0	22	AAW10367	Human oxy
16	76	100.0	22	AAW62536	Human CRH
17	76	100.0	22	AAW60848	Peptide p
18	76	100.0	30	AAW49299	19P2 liga
19	76	100.0	31	AAW1391	Human typ
20	76	100.0	31	AAW87615	Human 19P
21	76	100.0	31	AAW97235	Human typ
22	76	100.0	31	AAW49291	19P2 liga
23	76	100.0	31	AAW10362	Human oxy
24	76	100.0	31	AAW90991	Prolactin
25	76	100.0	31	AAW90995	Prolactin

26	76	100.0	31	4	AAW62531	Human CRH
27	76	100.0	31	5	AAW6401	Human PrR
28	76	100.0	31	6	ABU60843	Peptide p
29	76	100.0	31	6	ABU60827	Peptide p
30	76	100.0	31	7	ADC71228	Human pep
31	76	100.0	32	2	AAW31392	Human typ
32	76	100.0	32	3	AAW10363	Human oxy
33	76	100.0	32	4	AAW62532	Human CRH
34	76	100.0	32	6	ABU60844	Peptide p
35	76	100.0	33	2	AAW31393	Human typ
36	76	100.0	33	3	AAW10364	Human oxy
37	76	100.0	33	4	AAW62533	Human CRH
38	76	100.0	33	6	ABU60845	Peptide p
39	76	100.0	87	2	AAW31390	Human typ
40	76	100.0	87	2	AAW97226	Human typ
41	76	100.0	87	3	AAW10361	Human oxy
42	76	100.0	87	4	AAW62530	Human CRH
43	72	94.7	19	2	AAW31370	Bovine G
44	72	94.7	19	2	AAW95185	Bovine p
45	72	94.7	19	6	ABU60830	Peptide p

ALIGNMENTS

RESULT 1
AA49300
ID AA49300 standard; peptide; 14 AA.
XX
AC AA49300;
XX
DT 22-FEB-2000 (first entry)
XX
DE 19P2 ligand peptide fragment.
XX
KW Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;
KW pituitary; regulatory mechanism; central nervous system; pancreatic.
XX
OS Homo sapiens.
XX
FH Key
FT Modified-site 14
FT /note= "C-terminal amide"
XX
FN WO9960112-A1.
XX
PD 25-NOV-1999.
XX
PF 20-MAY-1999; 99WO-JP002650.
XX
XX 21-MAY-1999; 99JP-00140293.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Matsumoto H, Kitada C, Hinuma S;
XX
WPI; 2000-039381/03.
XX
New monoclonal antibodies, useful in diagnosis, as drugs and in studying diseases related to ligand abnormality.
XX
Disclosure; Page 27; 73pp; Japanese.

The invention provides a monoclonal antibody which has a specific reaction with the part peptide of the C-terminal of 19P2 ligand or its derivative. The antibodies can be used in diagnosis or to treat or prevent diseases associated with abnormality in the pituitary function regulatory mechanism (e.g. promotion of prolactin secretion), central nervous regulatory mechanism, and pancreatic function regulatory mechanism. The antibody-based immunoassay can also be applied in clarifying the physiological functions of the ligand and its derivative. Sequences AA49290-302 represent peptide fragments of the 19P2 ligand

SQ Sequence 14 AA;

Query Match 100.0%; Score 76; DB 3; Length 14;
 Best Local Similarity 100.0%; Pred. No. 2.2e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRG 13
 |||||
 DB 1 TPDINPAWYASRG 13

RESULT 2

AAW31394
 ID AAW31394 standard; peptide; 20 AA.

XX AAW31394;

DT 06-APR-1998 (first entry)

XX Human type G protein-coupled receptor ligand fragment 4.

XX G protein-coupled receptor; ligand binding; pharmaceutical; modulator;
 XX pituitary; central nervous system; pancreas; prophylactic;
 XX therapeutic agent.

OS Homo sapiens.

XX WC9724436-A2.

PN 10-JUL-1997.

PD 26-DEC-1996; 96WO-JF003821.

XX 28-DEC-1995; 95JP-00343371.

PR 15-MAR-1996; 96JP-00059419.

PR 12-AUG-1996; 96JP-00211805.

PR 18-SEP-1996; 96JP-00246573.

XX (TAKE) TAKEDA CHEM IND LTD.

PA Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;

PI Kitada C;

DR WPI; 1997-363672/33.

DR N-PSDB; AAW02431.

XX Claim 2; Page 185; 258pp; English.

XX This sequence represents a peptide fragment from a novel human type
 CC ligand polypeptide corresponding to amino acid residues 34 to 53 of the
 CC sequence represented in AAW31390 and is used in an assay to monitor
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical
 CC compositions containing this ligand may be used as a pituitary function
 CC modulator, a central nervous system modulator or a pancreatic function
 CC modulator. This ligand could have specific applications as a prophylactic
 CC or therapeutic agent for dementia, depression, hyperkinetic syndrome,
 CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,
 CC growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,
 CC hypercholesterolaemia, hypoglycaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,
 CC amyotrophic lateral sclerosis, acute myocardial infarction, infertility,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis and/or oligogalactia. Assays can also be developed to screen
 CC compounds which are capable of altering the binding activity of the
 CC ligand affecting activation of the G protein-coupled receptor protein

SQ Sequence 20 AA;

Query Match

Best Local Similarity 100.0%; Score 76; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 3.3e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRG 13

DB 1 TPDINPAWYASRG 13

RESULT 3

AAW97236

ID AAW97236 standard; peptide; 20 AA.

XX AAW97236;

DT 06-MAY-1999 (first entry)

XX Human type ligand polypeptide fragment.

XX Rat type ligand; modulation; prolactin secretion;
 KW G protein-coupled receptor; gpCR; hypovarianism; gonocyst cocogenesis;
 KW menopausal syndrome; euthyroid; hypometabolism; lactation;
 KW pituitary adenomatosis; brain tumour; amenorrhoea; autoimmune disease;
 KW prolactinoma; infertility; impotence; amenorrhoea; galactorrhea;
 KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;
 KW contraceptive; placental function; choriocarcinoma; hydatid mole;
 KW abortion; irruption mole; abnormal saccharometabolism;
 KW abnormal lipidmetabolism; oxytocia.

XX Homo sapiens.

XX WO9858962-A1.

XX 30-DEC-1998.

XX 22-JUN-1998; 98WO-JF002765.

XX 23-JUN-1997; 97JP-00165437.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Hinuma S, Kawamata Y, Fujii R, Matsumoto H;

XX WPI; 1999-105614/09.

XX Use of G protein-coupled receptor ligands - for modulating prolactin
 PT secretion or placental function, e.g. for treating menopausal syndrome,
 PT tumours, autoimmune disease or abnormal pregnancy.

XX Claim 3; Page 166; 241pp; English.

XX The present sequence represents a human type ligand fragment. It is used
 CC in the course of the invention. The specification describes an agent for
 CC modulating prolactin secretion which comprises a ligand polypeptide or a
 CC salt, for a G protein-coupled receptor (gpCR) protein. The agents for
 CC promoting prolactin secretion can be used for treating or preventing
 CC hypovarianism, gonocyst cocogenesis, menopausal syndrome, euthyroid or
 CC hypometabolism. They can be used for promoting lactation in a domestic
 CC mammal and as an aphrodisiac. The agents for inhibiting prolactin
 CC secretion can be used for treating or preventing pituitary adenomatosis,
 CC brain tumour, amenorrhoea, autoimmune disease, prolactinoma, Chiari-
 CC infertile, impotence, amenorrhoea, galactorrhea, acromegaly, Chiari-
 CC Frommel syndrome, Argonz-del Castillo syndrome, Forbes-Albright syndrome,
 CC lymphoma, Sheehan syndrome or dyszoospermia. The inhibitory agents can
 CC also be used as contraceptives. The agents for modulating placental
 CC function can be used for treating or preventing choriocarcinoma, hydatid
 CC mole, irruption mole, abortion, unthrifty fetus, abnormal
 CC saccharometabolism, abnormal lipidmetabolism or oxytocia

XX Sequence 20 AA;

Query Match

Best Local Similarity 100.0%; Score 76; DB 2; Length 20;
 Pred. No. 3.3e-06;

XX Human oxytocin secretion promoting peptide SEQ ID NO: 35.
DE
XX Human; oxytocin secretion promoter; G protein-coupled receptor protein;
KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
KW caesarean section; artificial fertilization; galactostasis; goat; pig;
KW veterinary medicine; milk production.
XX
XX Homo sapiens.
XX WO200038704-A1.
XX
XX 06-JUL-2000.
XX
XX 22-DEC-1999; 99WO-JP007199.
XX
XX 25-DEC-1998; 98P-00369585.
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Matsumoto H, Kitada C, Hinuma S;
XX
XX WPI; 2000-452298/39.
XX
XX Physiologically-active polypeptide recognized as ligand by G protein-
PT coupled receptor protein, for promoting secretion of oxytocin, as drugs
PT for diseases relating to oxytocin secretion and in veterinary medicine.
XX
XX Disclosure; Page 63; 72pp; Japanese.
XX
XX This invention describes a novel oxytocin secretion-regulating agent
CC which contains a ligand peptide or its salt for the G protein-coupled
CC receptor protein. It is useful in the form of drugs for ameliorating,
CC preventing and treating diseases relating to oxytocin secretion e.g. weak
CC pains and atonic bleeding, before and after expulsion of placenta,
CC uterine recovery failure, caesarean section, stoppage of artificial
CC fertilization or galactostasis and is also applicable in veterinary
CC medicine for promoting milk production in cow, goat and pig. This
CC sequence represents a human peptide which acts as an oxytocin secretion
CC promoter
XX
XX Sequence 20 AA;
SQ
Query Match 100.0%; Score 76; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TPDINPAWYASRG 13
DB 1 TPDINPAWYASRG 13
RESULT 6
AAB90992
ID AAB90992 standard; peptide; 20 AA.
XX
XX AAB90992;
XX
XX 22-JUN-2001 (first entry)
DT
XX
XX Prolactin releasing peptide SEQ ID NO:166.
DE
XX
XX Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimidy; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
XX Homo sapiens.
OS
OS Synthetic.
XX
XX WO200069900-A2.
FN
XX
XX 23-NOV-2000.
PD
XX

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TPDINPAWYASRG 13
DB 1 TPDINPAWYASRG 13
RESULT 4
AA49294
ID AA49294 standard; peptide; 20 AA.
XX
XX AC AA49294;
XX
XX 22-FEB-2000 (first entry)
DT
XX
XX 19P2 ligand peptide fragment.
DE
XX
XX Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;
KW pituitary; regulatory mechanism; central nervous system; pancreatic.
KW
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Modified-site 20
FT /note= "C-terminal amide"
XX
XX WO960112-A1.
FN
XX
XX 25-NOV-1999.
PD
XX
XX 20-MAY-1999; 99WO-JP002650.
PF
XX
XX 21-MAY-1998; 98JP-00140293.
PR
XX
XX (TAKE) TAKEDA CHEM IND LTD.
PA
XX
XX Matsumoto H, Kitada C, Hinuma S;
PI
XX
XX WPI; 2000-039381/03.
DR
XX
XX New monoclonal antibodies, useful in diagnosis, as drugs and in studying
PT diseases related to ligand abnormality.
PT
XX
XX Disclosure; Page 26; 73pp; Japanese.
PS
XX
XX The invention provides a monoclonal antibody which has a specific
CC reaction with the part peptide of the C-terminal of 19P2 ligand or its
CC derivative. The antibodies can be used in diagnosis or to treat or
CC prevent diseases associated with abnormality in the pituitary function
CC regulatory mechanism (e.g. promotion of prolactin secretion), central
CC nervous regulatory mechanism, and pancreatic function regulatory
CC mechanism. The antibody-based immunoassay can also be applied in
CC clarifying the physiological functions of the ligand and its derivative.
CC Sequences AA49290-302 represent peptide fragments of the 19P2 ligand
XX
XX
XX Sequence 20 AA;
SQ
Query Match 100.0%; Score 76; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TPDINPAWYASRG 13
DB 1 TPDINPAWYASRG 13
RESULT 5
AAB10365
ID AAB10365 standard; peptide; 20 AA.
XX
XX AC AAB10365;
XX
XX 24-NOV-2000 (first entry)
DT

PF 17-MAY-2000; 2000WO-US013576.
 XX 17-MAY-1999; 99US-0134406P.
 PR 10-SEP-1999; 99US-0153406P.
 PR 15-OCT-1999; 99US-0159783P.
 XX (CONJ-) CONJUCHEM INC.
 PA Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
 PI WPI; 2001-112059/12.
 XX
 XX Modifying and attaching therapeutic peptides to albumin prevents
 PT peptidase degradation, useful for increasing length of in vivo activity.
 XX
 XX Disclosure; Page 244; 733pp; English.
 XX The present invention describes a modified therapeutic peptide (I)
 CC comprising a therapeutically active amino acid region (III) and a
 CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a
 CC peptide stabilised therapeutic peptide composed of 3-50 amino acids.
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
 CC factors and neurotransmitters, to protect them from peptidase activity in
 CC vivo for the treatment of various disorders. Endogenous therapeutic
 CC peptides are not suitable as drug candidates as they require frequent
 CC administration due to rapid degradation by peptidases in the body.
 CC Modifying and attaching therapeutic peptides to albumin prevents or
 CC reduces the action of peptidases to increase length of activity (half
 CC life) and specificity as bonding to large molecules decreases
 CC intracellular uptake and interference with physiological processes.
 CC AAB90829 to AAB92441 represent peptides which can be used in the
 CC exemplification of the present invention
 XX
 XX SQ Sequence 20 AA;
 Query Match 100.0%; Score 76; DB 4; Length 20;
 Best Local Similarity 100.0%; Pred. No. 3.3e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TPDINPAWYASRG 13
 |||||
 Db 1 TPDINPAWYASRG 13
 |||||
 RESULT 7
 AAG62534
 ID AAG62534 standard; peptide; 20 AA.
 XX
 XX AAG62534;
 DT 24-AUG-2001 (first entry)
 XX
 DE Human CRH releasing protein related peptide SEQ ID NO: 35.
 XX
 XX Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;
 KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;
 KW Addison's disease; adrenal gland hyperfunction; obesity.
 OS Homo sapiens.
 OS
 XX WO200135984-A1.
 PN
 XX 25-MAY-2001.
 PD
 XX
 XX 17-NOV-2000; 2000WO-JP008119.
 PF
 PR 18-NOV-1999; 99JP-00327900.
 PR 26-SEP-2000; 2000JP-00297073.
 XX
 XX (TAKE) TAKEDA CHEM IND LTD.
 PA

PI Kitada C, Matsumoto H, Hinuma S;
 XX WPI; 2001-355552/37.
 XX
 PT Use of G protein receptor ligand or peptide for controlling corticotropin
 PT releasing hormone secretion.
 XX
 PS Claim 4; Page 75; 90pp; Japanese.
 XX
 CC The present sequence describes a method of controlling the secretion of
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein
 CC receptor ligand. This can be used to control the secretion of CRH and is
 CC useful as an analgesic or for treating, preventing or ameliorating
 CC diseases associated with CRH secretion such as hyperaldosteronism,
 CC hypercortisolaemia, secondary or chronic hypoadrenocorticism, Addison's
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The
 CC present sequence is a peptide used in the exemplification of the
 CC invention
 XX
 XX SQ Sequence 20 AA;
 Query Match 100.0%; Score 76; DB 4; Length 20;
 Best Local Similarity 100.0%; Pred. No. 3.3e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TPDINPAWYASRG 13
 |||||
 Db 1 TPDINPAWYASRG 13
 |||||
 RESULT 8
 AAE26404
 ID AAE26404 standard; peptide; 20 AA.
 XX
 XX AAE26404;
 DT 13-DEC-2002 (first entry)
 XX
 DE Human PRP-31 C-terminal peptide, PRP-20.
 XX
 KW Human; wakefulness; sleep disorder; prolactin releasing peptide receptor;
 KW PRP; GPR10; therapy; epilepsy; narcolepsy; sleepiness; sleep apnoea;
 KW insomnia; idiopathic hypersomnia; psychogenic hypersomnia; seizure;
 KW anticonvulsant.
 OS Homo sapiens.
 OS
 XX US2002037533-A1.
 PN
 XX 28-MAR-2002.
 PD
 XX 17-AUG-2001; 2001US-00932161.
 PF
 PR 28-APR-2000; 2000US-00560915.
 XX
 XX (CIVE/) CIVELLI O.
 PA (LINS/) LIN S.
 XX
 XX Civelli O, Lin S;
 PI
 XX WPI; 2002-403931/43.
 DR
 XX Screening for compounds useful for promoting wakefulness or sleep, and
 PT for treating sleeping disorders, e.g. insomnia, hypersomnia or sleep
 PT apnea, comprises administering a prolactin releasing peptide agonist or
 PT antagonist.
 XX
 PS Disclosure; Page 25; 35pp; English.
 XX
 CC The present invention relates to a method of screening for compounds for
 CC promoting wakefulness or sleep in a mammal. The method involves
 CC administering a prolactin releasing peptide (PrRP) receptor (GPR10)

agonist or antagonist respectively and determining the ability of the compound to promote wakefulness or sleep. The compounds identified from the method are used in the therapy of epilepsy and other diseases associated with absence seizures and in promoting wakefulness and sleep in individuals having sleep disorders such as insomnia and narcolepsy. PRP receptor agonists may be used to treat common disorders which lead to sleepiness, e.g. sleep apnoea, narcolepsy, idiopathic hypersomnia and psychogenic hypersomnia. PRP receptor antagonists are useful for promoting sleep and for treating insomnia such as adjustment sleep disorder and psychophysiological insomnia. The present sequence is human PRP-31 C-terminal peptide, PrRP-20

Sequence 20 AA;

Query Match 100.0%; Score 76; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAMYASRG 13
|||||
DB 1 TPDINPAMYASRG 13

RESULT 9
ABU60846
ID ABU60846 standard; peptide; 20 AA.
XX
AC ABU60846;
XX
XX 06-MAY-2003 (first entry)
XX
DE Peptide production by gene recombination associated peptide #30.
XX
XX Peptide production; low-molecular peptide; KiSS-1; GPR8 ligand;
KW gene recombination.
XX
XX Homo sapiens.
XX
XX WO200292829-A1.
XX
XX 21-NOV-2002.
XX
XX 16-MAY-2002; 2002WO-JP004735.
XX
XX 17-MAY-2001; 2001JP-00147341.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Nishimura O, Suenaga M, Ito T, Kitada C;
XX
XX WPI; 2003-129302/12.
XX

Process for producing peptides e.g. KiSS-1 peptide and GPR8 ligand for subsequent applications by gene recombination technique through tandem repeats to provide precursor protein with specific cleavage sites.

Disclosure; Page 68; 87pp; Japanese.

The invention describes a method of producing a peptide comprising the excision of the N and C-terminals of a target peptide with enzymes or chemically through the attached cleavage sites repeated by ligation in a precursor protein. The method is for producing (low-molecular) peptides e.g. KiSS-1 peptide and GPR8 ligand for subsequent applications by the gene recombination technique through tandem repeats to provide a precursor protein with specific cleavage sites. With this method, peptide production can be carried out easily to provide large quantities of the required peptides. This is the amino acid sequence of a peptide associated with the peptide production method of the invention

Sequence 20 AA;

Query Match 100.0%; Score 76; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAMYASRG 13
|||||
DB 1 TPDINPAMYASRG 13

RESULT 10
AAW31395
ID AAW31395 standard; peptide; 21 AA.
XX
AC AAW31395;
XX
DT 06-APR-1998 (first entry)
XX
XX Human type G protein-coupled receptor ligand fragment 5.
DE
XX
XX G protein-coupled receptor; ligand binding; pharmaceutical; modulator;
KW pituitary; central nervous system; pancreas; prophylactic;
KW therapeutic agent.
XX
XX Homo sapiens.
OS
XX WO9724436-A2.
FN
XX 10-JUL-1997.
PD
XX 26-DEC-1996; 96WO-JP003821.
XX
XX 28-DEC-1995; 95JP-00343371.
PR 15-MAR-1996; 96JP-00059419.
PR 12-AUG-1996; 96JP-00211805.
PR 18-SEP-1996; 96JP-00246573.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
PA
XX
XX Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;
PI Kitada C;
XX
XX WPI; 1997-363672/33.
DR N-PSDB; AAV02432.
XX
XX Ligand peptide for G protein-coupled receptor - acts by modulating
PT function in the central nervous system, pancreas and pituitary gland.
XX
XX Claim 2; Page 186; 258pp; English.

This sequence represents a peptide fragment from a novel human type ligand polypeptide corresponding to amino acid residues 34 to 54 of the sequence represented in AAW31395 and is used in an assay to monitor ligand binding to the G protein-coupled receptor protein. Pharmaceutical compositions containing this ligand may be used as a pituitary function modulator, a central nervous system modulator or a pancreatic function modulator. This ligand could have specific applications as a prophylactic or therapeutic agent for dementia, depression, hyperkinetic syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone secretory disease, hyper- and polynephria, hyperlipidaemia, hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis, acute myocardial infarction, infertility, osteoporosis and/or oligosaccharia. Assays can also be developed to screen compounds which are capable of altering the binding activity of the ligand affecting activation of the G protein-coupled receptor protein

Sequence 21 AA;

Query Match 100.0%; Score 76; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAMYASRG 13

Db 1 TPDPINPAWASRG 13
 |||||
 RESULT 11
 AAB10366
 ID AAB10366 standard; peptide; 21 AA.
 XX AC AAB10366;
 XX DT 24-NOV-2000 (first entry)
 XX DE Human oxytocin secretion promoting peptide SEQ ID NO: 36.
 XX KW Human, oxytocin secretion promoter; G protein-coupled receptor protein;
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;
 KW veterinary medicine; milk production.
 XX OS Homo sapiens.
 XX PN WO200038704-A1.
 XX XX 06-JUL-2000.
 XX PF 22-DEC-1999; 99WO-JP007199.
 XX PR 25-DEC-1998; 98JP-00369585.
 XX PA (TAKE) TAKEDA CHEM IND LTD.
 XX PI Matsumoto H, Kitada C, Hinuma S;
 XX WPI; 2000-452298/39.
 XX PT Physiologically-active polypeptide recognized as ligand by G protein-
 PT coupled receptor protein, for promoting secretion of oxytocin, as drugs
 PT for diseases relating to oxytocin secretion and in veterinary medicine.
 XX PS Disclosure; Page 63; 72pp; Japanese.
 XX CC This invention describes a novel oxytocin secretion-regulating agent
 CC which contains a ligand peptide or its salt for the G protein-coupled
 CC receptor protein. It is useful in the form of drugs for ameliorating,
 CC preventing and treating diseases relating to oxytocin secretion e.g. weak
 CC pains and atonic bleeding, before and after expulsion of placenta,
 CC uterine recovery failure, caesarean section, stoppage of artificial
 CC fertilization or galactostasis and is also applicable in veterinary
 CC medicine for promoting milk production in cow, goat and pig. This
 CC sequence represents a human peptide which acts as an oxytocin secretion
 CC promoter
 XX SQ Sequence 21 AA;
 Query Match 100.0%; Score 76; DB 3; Length 21;
 Best Local Similarity 100.0%; Pred. No. 3.5e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 TPDPINPAWASRG 13
 |||||
 Db 1 TPDPINPAWASRG 13
 |||||
 RESULT 12
 AAG62535
 ID AAG62535 standard; peptide; 21 AA.
 XX AC AAG62535;
 XX DT 24-AUG-2001 (first entry)
 XX DE Human CRH releasing protein related peptide SEQ ID NO: 36.
 XX XX

KW Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;
 KW analgesic; hyperaldosteronism; hypercortisolism; hypoadrenocorticism;
 KW Addison's disease; adrenal gland hyperfunction; obesity.
 XX OS Homo sapiens.
 XX PN WO200135984-A1.
 XX PD 25-MAY-2001.
 XX PF 17-NOV-2000; 2000WO-JP008119.
 XX PR 18-NOV-1999; 99JP-00327900.
 XX PR 26-SEP-2000; 2000JP-00297073.
 XX PA (TAKE) TAKEDA CHEM IND LTD.
 XX PI Kitada C, Matsumoto H, Hinuma S;
 XX WPI; 2001-355552/37.
 XX XX Use of G protein receptor ligand or peptide for controlling corticotropin
 PT releasing hormone secretion.
 XX PS Disclosure; Page 75; 90pp; Japanese.
 XX CC The present sequence describes a method of controlling the secretion of
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein
 CC receptor ligand. This can be used to control the secretion of CRH and is
 CC useful as an analgesic or for treating, preventing or ameliorating
 CC diseases associated with CRH secretion such as hyperaldosteronism,
 CC hypercortisolism, secondary or chronic hypoadrenocorticism, Addison's
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The
 CC present sequence is a peptide used in the exemplification of the
 CC invention
 XX SQ Sequence 21 AA;
 Query Match 100.0%; Score 76; DB 4; Length 21;
 Best Local Similarity 100.0%; Pred. No. 3.5e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 TPDPINPAWASRG 13
 |||||
 Db 1 TPDPINPAWASRG 13
 |||||
 RESULT 13
 ABU60847
 ID ABU60847 standard; peptide; 21 AA.
 XX AC ABU60847;
 XX DT 06-MAY-2003 (first entry)
 XX DE Peptide production by gene recombination associated peptide #31.
 XX KW Peptide production; low-molecular peptide; KiSS-1; GPR8 ligand;
 KW gene recombination.
 XX OS Homo sapiens.
 XX PN WO200292829-A1.
 XX PD 21-NOV-2002.
 XX PF 16-MAY-2002; 2002WO-JP004735.
 XX PR 17-MAY-2001; 2001JP-00147341.
 XX PA (TAKE) TAKEDA CHEM IND LTD.
 XX XX

ligand polypeptide corresponding to amino acid residues 34 to 53 of the sequence represented in AAW31390 and is used in an assay to monitor ligand binding to the G protein-coupled receptor protein. Pharmaceutical compositions containing this ligand may be used as a pituitary function modulator, a central nervous system modulator or a pancreatic function modulator. This ligand could have specific applications as a prophylactic or therapeutic agent for dementia, depression, hyperkinetic syndrome, or disturbance of consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia, hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis, acute myocardial infarction, infertility, spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis, osteoporosis and/or oligogalactaemia. Assays can also be developed to screen compounds which are capable of altering the binding activity of the ligand affecting activation of the G protein-coupled receptor protein

Sequence 22 AA;

Query Match 100.0%; Score 76; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 3.7e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWVASRG 13
DQ 1 TPDINPAWVASRG 13
|||||

RESULT 15
AAB10367
ID AAB10367 standard; peptide; 22 AA.
XX AAB10367;
AC
XX
XX
XX
DT 24-NOV-2000 (first entry)
XX
DE Human oxytocin secretion promoting peptide SEQ ID NO: 37.
XX
XX Human; oxytocin secretion promoter; G protein-coupled receptor protein;
KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
KW caesarean section; artificial fertilization; galactostasis; goat; pig;
XX veterinary medicine; milk production.
XX
XX Homo sapiens.
OS
XX
XX WO200038704-A1.
PN
XX
XX
PD 06-JUL-2000.
XX
XX 22-DEC-1999; 99WO-JP007199.
PF
XX 25-DEC-1998; 98JP-00369585.
PR
XX
XX (TAKE) TAKEDA CHEM IND LTD.
PA
XX
XX Matsumoto H, Kitada C, Hinuma S;
FI
XX
XX WPT; 2000-452298/39.
DR
XX
XX Physiologically-active polypeptide recognized as ligand by G protein-
PT coupled receptor protein, for promoting secretion of oxytocin, as drugs
PT for diseases relating to oxytocin secretion and in veterinary medicine.
XX
XX Disclosure; Page 64; 72pp; Japanese.
PS
XX This invention describes a novel oxytocin secretion-regulating agent
CC which contains a ligand peptide or its salt for the G protein-coupled
CC receptor protein. It is useful in the form of drugs for ameliorating,
CC preventing and treating diseases relating to oxytocin secretion e.g. weak
CC pains and atonic bleeding, before and after expulsion of placenta,
CC uterine recovery failure, caesarean section, stoppage of artificial
CC fertilization or galactostasis and is also applicable in veterinary

CC medicine for promoting milk production in cow, goat and pig. This
 CC sequence represents a human peptide which acts as an oxytocin secretion
 CC promoter
 XX
 SQ Sequence 22 AA;

Query Match 100.0%; Score 76; DB 3; Length 22;
 Best Local Similarity 100.0%; Pred. NO. 3.7e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRG 13
 |||||
 Db 1 TPDINPAWYASRG 13

Search completed: August 12, 2004, 14:43:54
 Job time : 27.7122 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:49:10 ; Search time 22.9767 Seconds
(without alignments)
177.617 Million cell updates/sec

Title: US-09-700-643A-2_COPY_12_24
Perfect score: 76
Sequence: 1 TPDINPAWYASRG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1292805 seqs, 313927144 residues

Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCR_NEW PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCR05_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	100.0	20	9	US-09-932-161-18
2	76	100.0	20	14	US-10-044-592-18
3	76	100.0	31	9	US-09-932-161-15
4	76	100.0	31	14	US-10-044-592-15
5	76	100.0	87	13	US-10-044-592-92
6	72	94.7	19	13	US-10-044-592-27
7	72	94.7	20	9	US-09-932-161-16
8	72	94.7	20	13	US-10-044-592-42
9	72	94.7	20	14	US-10-044-592-16
10	72	94.7	21	13	US-10-044-592-43
11	72	94.7	22	13	US-10-044-592-44
12	72	94.7	25	13	US-10-044-592-78
13	72	94.7	29	13	US-10-044-592-26
14	72	94.7	31	9	US-09-932-161-13
15	72	94.7	31	13	US-10-044-592-39

Sequence 13, Appl
Sequence 40, Appl
Sequence 41, Appl
Sequence 28, Appl
Sequence 38, Appl
Sequence 82, Appl
Sequence 84, Appl
Sequence 86, Appl
Sequence 88, Appl
Sequence 17, Appl
Sequence 6, Appl
Sequence 17, Appl
Sequence 14, Appl
Sequence 4, Appl
Sequence 5, Appl
Sequence 14, Appl
Sequence 90, Appl
Sequence 1, Appl
Sequence 96, Appl
Sequence 94, Appl
Sequence 8, Appl
Sequence 197, Appl
Sequence 198301,
Sequence 50882, A
Sequence 6327, A
Sequence 1094, A
Sequence 1094, A
Sequence 248980,
Sequence 268092,
Sequence 469, Appl

16 72 94.7 31 14 US-10-096-777-13
17 72 94.7 32 13 US-10-044-592-40
18 72 94.7 33 13 US-10-044-592-41
19 72 94.7 38 13 US-10-044-592-28
20 72 94.7 98 13 US-10-044-592-38
21 72 94.7 98 13 US-10-044-592-82
22 72 94.7 98 13 US-10-044-592-84
23 72 94.7 98 13 US-10-044-592-86
24 72 94.7 98 13 US-10-044-592-88
25 68 89.5 20 9 US-09-932-161-17
26 68 89.5 20 13 US-10-044-592-6
27 68 89.5 20 14 US-10-096-777-17
28 68 89.5 31 9 US-09-932-161-14
29 68 89.5 31 13 US-10-044-592-4
30 68 89.5 31 13 US-10-044-592-5
31 68 89.5 31 14 US-10-096-777-14
32 68 89.5 70 13 US-10-044-592-90
33 68 89.5 82 13 US-10-044-592-1
34 68 89.5 86 13 US-10-044-592-96
35 68 89.5 91 13 US-10-044-592-94
36 57 75.0 9 13 US-10-044-592-8
37 47 61.8 465 14 US-10-301-822-197
38 45 59.2 816 16 US-10-437-963-198301
39 44 57.9 387 12 US-10-282-122A-50882
40 44 57.9 527 15 US-10-369-493-6327
41 43 56.6 44 9 US-09-925-299-1094
42 43 56.6 44 10 US-09-925-299-1094
43 43 56.6 59 12 US-10-424-599-248980
44 43 56.6 183 12 US-10-424-599-268092
45 42.5 55.9 664 12 US-10-389-647-469

ALIGNMENTS

RESULT 1

US-09-932-161-18
; Sequence 18, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Civealli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-932-161-18

Query Match 100.0% Score 76; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRG 13

Db 1 TPDINPAWYASRG 13

RESULT 2

US-10-096-777-18
; Sequence 18, Application US/10096777
; Publication No. US20030171270A1
; GENERAL INFORMATION:
; APPLICANT: Civealli, Olivier
; APPLICANT: Lin, Steven

```

; TITLE OF INVENTION: Therapeutic Compositions and Methods
; FILE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-096-777-18

```

```

Query Match      100.0%; Score 76; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 TPDINPAWYASRG 13
      |||||
Db      1 TPDINPAWYASRG 13

```

RESULT 3

```

US-09-932-161-15
; Sequence 15, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Civealli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; FILE OF INVENTION: Promoting Wakefulness and Sleep
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US/09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-932-161-15

```

```

Query Match      100.0%; Score 76; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. No. 5.1e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 TPDINPAWYASRG 13
      |||||
Db      12 TPDINPAWYASRG 24

```

RESULT 4

```

US-10-096-777-15
; Sequence 15, Application US/10096777
; Publication No. US20030171270A1
; GENERAL INFORMATION:
; APPLICANT: Civealli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; FILE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 31

```

```

; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-096-777-15

Query Match      100.0%; Score 76; DB 14; Length 31;
Best Local Similarity 100.0%; Pred. No. 5.1e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 TPDINPAWYASRG 13
      |||||
Db      12 TPDINPAWYASRG 24

```

RESULT 5

```

US-10-044-592-92
; Sequence 92, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 92
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-044-592-92

```

```

Query Match      100.0%; Score 76; DB 13; Length 87;
Best Local Similarity 100.0%; Pred. No. 0.00014;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 TPDINPAWYASRG 13
      |||||
Db      34 TPDINPAWYASRG 46

```

RESULT 6

```

US-10-044-592-27
; Sequence 27, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 27
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-27

```

Query Match

94.7%; Score 72; DB 13; Length 19;

Best Local Similarity 92.3%; Pred. No. 0.00014;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TPDINPAWYASRG 13
Db 1 TPDINPAWYASRG 13

RESULT 7

US-09-932-161-16
; Sequence 16, Application US/09932161
; Patent No. US20020037533A1

GENERAL INFORMATION:
; APPLICANT: Civelletti, Olivier

APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; TITLE OF INVENTION: Promoting Wakefulness and Sleep

FILE REFERENCE: P-UC 4679

CURRENT APPLICATION NUMBER: US/09/932,161

PRIOR FILING DATE: 2001-08-17

PRIOR APPLICATION NUMBER: US 09/560,915

PRIOR FILING DATE: 2000-04-28

NUMBER OF SEQ ID NOS: 24

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 16

LENGTH: 20

TYPE: PRT

ORGANISM: Bos taurus

US-09-932-161-16

Query Match 94.7%; Score 72; DB 9; Length 20;

Best Local Similarity 92.3%; Pred. No. 0.00015;

Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TPDINPAWYASRG 13
Db 1 TPDINPAWYASRG 13

RESULT 8

US-10-044-592-42
; Sequence 42, Application US/10044592
; Publication No. US20020143152A1

GENERAL INFORMATION:
; APPLICANT: Hiruma, Shuji

APPLICANT: Fukusumi, Shoji

FILE REFERENCE: 2463US2P

CURRENT APPLICATION NUMBER: US/10/044,592

PRIOR FILING DATE: 2002-01-10

PRIOR APPLICATION NUMBER: US 09/403639

PRIOR FILING DATE: 1999-25-10

PRIOR APPLICATION NUMBER: PCT/JF98/01923

PRIOR FILING DATE: 1998-04-27

NUMBER OF SEQ ID NOS: 96

SOFTWARE:

SEQ ID NO 42

LENGTH: 20

TYPE: PRT

ORGANISM: Bovine

US-10-044-592-42

Query Match 94.7%; Score 72; DB 13; Length 20;

Best Local Similarity 92.3%; Pred. No. 0.00015;

Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TPDINPAWYASRG 13
Db 1 TPDINPAWYASRG 13

RESULT 9

US-10-096-777-16

; Sequence 16, Application US/10096777

GENERAL INFORMATION:
; APPLICANT: Civelletti, Olivier

APPLICANT: Lin, Steven

FILE REFERENCE: P-UC 3534

CURRENT APPLICATION NUMBER: US/10/096,777

PRIOR FILING DATE: 2002-03-12

PRIOR APPLICATION NUMBER: US/09/560,915

NUMBER OF SEQ ID NOS: 24

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 16

LENGTH: 20

TYPE: PRT

ORGANISM: Bos taurus

US-10-096-777-16

Query Match 94.7%; Score 72; DB 14; Length 20;

Best Local Similarity 92.3%; Pred. No. 0.00015;

Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TPDINPAWYASRG 13
Db 1 TPDINPAWYASRG 13

RESULT 10

US-10-044-592-43

; Sequence 43, Application US/10044592

GENERAL INFORMATION:
; APPLICANT: Hiruma, Shuji

APPLICANT: Fukusumi, Shoji

FILE REFERENCE: 2463US2P

CURRENT APPLICATION NUMBER: US/10/044,592

PRIOR FILING DATE: 2002-01-10

PRIOR APPLICATION NUMBER: PCT/JF98/01923

PRIOR FILING DATE: 1998-04-27

NUMBER OF SEQ ID NOS: 96

SOFTWARE:

SEQ ID NO 43

LENGTH: 21

TYPE: PRT

ORGANISM: Bovine

US-10-044-592-43

Query Match 94.7%; Score 72; DB 13; Length 21;

Best Local Similarity 92.3%; Pred. No. 0.00015;

Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TPDINPAWYASRG 13
Db 1 TPDINPAWYASRG 13

RESULT 11

US-10-044-592-44

; Sequence 44, Application US/10044592

GENERAL INFORMATION:
; APPLICANT: Hiruma, Shuji

APPLICANT: Fukusumi, Shoji

FILE REFERENCE: 2463US2P

CURRENT APPLICATION NUMBER: US/10/044,592

PRIOR FILING DATE: 2002-01-10

PRIOR APPLICATION NUMBER: PCT/JF98/01923

PRIOR FILING DATE: 1998-04-27

NUMBER OF SEQ ID NOS: 96

SOFTWARE:

SEQ ID NO 43

LENGTH: 21

TYPE: PRT

ORGANISM: Bovine

US-10-044-592-44

```

; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 44
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-44

```

```

Query Match          94.7%; Score 72; DB 13; Length 22;
Best Local Similarity 92.3%; Pred. No. 0.00016;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 TPDINPAWYASRG 13
| | | | | | | | | |
Db 1 TPDINPAWYAGRG 13

```

```

RESULT 12
US-10-044-592-78
; Sequence 78, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 78
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Bovine
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(20)
; OTHER INFORMATION: primer
; NAME/KEY: misc_feature
; LOCATION: (52)..(76)
; OTHER INFORMATION: primer
US-10-044-592-78

```

```

Query Match          94.7%; Score 72; DB 13; Length 25;
Best Local Similarity 92.3%; Pred. No. 0.00018;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 TPDINPAWYASRG 13
| | | | | | | | | |
Db 10 TPDINPAWYAGRG 22

```

```

RESULT 13
US-10-044-592-26
; Sequence 26, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji

```

```

; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 26
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-26

```

```

Query Match          94.7%; Score 72; DB 13; Length 29;
Best Local Similarity 92.3%; Pred. No. 0.00021;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 TPDINPAWYASRG 13
| | | | | | | | | |
Db 12 TPDINPAWYAGRG 24

```

```

RESULT 14
US-09-932-161-13
; Sequence 13, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Clivelli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-932-161-13

```

```

Query Match          94.7%; Score 72; DB 9; Length 31;
Best Local Similarity 92.3%; Pred. No. 0.00022;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 TPDINPAWYASRG 13
| | | | | | | | | |
Db 12 TPDINPAWYAGRG 24

```

```

RESULT 15
US-10-044-592-39
; Sequence 39, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923

```

Thu Aug 12 15:25:16 2004

```

; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 39
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-39

Query Match      94.7%; Score 72; DB 13; Length 31;
Best/Local Similarity 92.3%; Pred. No. 0.00022;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 TPDINPAWYASRG 13
      |||||
Db      12 TPDINPAWYASRG 24

Search completed: August 12, 2004, 15:22:50
Job time : 22.9767 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:36 ; Search time 7.40698 Seconds
(without alignments)
90.609 Million cell updates/sec

Title: US-09-700-643A-2_COPY_12_24

Perfect score: 76

Sequence: 1 TPDINPAWASRG 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/prodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/prodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/prodata/2/iaa/PTUS_COMB.pep.*
- 6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	20	3 US-09-105-678A-46	Sequence 46, Appl
2	76	100.0	20	3 US-08-776-971-64	Sequence 64, Appl
3	76	100.0	20	3 US-09-421-208-46	Sequence 46, Appl
4	76	100.0	20	4 US-09-560-915-18	Sequence 18, Appl
5	76	100.0	21	3 US-09-105-678A-47	Sequence 47, Appl
6	76	100.0	21	3 US-08-776-971-65	Sequence 65, Appl
7	76	100.0	21	3 US-09-421-208-47	Sequence 47, Appl
8	76	100.0	22	3 US-09-105-678A-48	Sequence 48, Appl
9	76	100.0	22	3 US-08-776-971-66	Sequence 66, Appl
10	76	100.0	22	3 US-09-421-208-48	Sequence 48, Appl
11	76	100.0	31	3 US-09-105-678A-9	Sequence 9, Appl
12	76	100.0	31	3 US-09-105-678A-43	Sequence 43, Appl
13	76	100.0	31	3 US-08-776-971-61	Sequence 61, Appl
14	76	100.0	31	3 US-09-421-208-9	Sequence 9, Appl
15	76	100.0	31	3 US-09-421-208-43	Sequence 43, Appl
16	76	100.0	31	4 US-09-560-915-15	Sequence 15, Appl
17	76	100.0	32	3 US-09-105-678A-44	Sequence 44, Appl
18	76	100.0	32	3 US-08-776-971-62	Sequence 62, Appl
19	76	100.0	32	3 US-09-421-208-44	Sequence 44, Appl
20	76	100.0	33	3 US-09-105-678A-45	Sequence 45, Appl
21	76	100.0	33	3 US-08-776-971-63	Sequence 63, Appl
22	76	100.0	33	3 US-09-421-208-45	Sequence 45, Appl
23	76	100.0	87	3 US-08-776-971-59	Sequence 59, Appl
24	76	100.0	87	3 US-08-776-971-135	Sequence 135, App
25	76	100.0	87	3 US-08-776-971-138	Sequence 138, App
26	72	94.7	19	3 US-09-105-678A-30	Sequence 30, Appl
27	72	94.7	19	3 US-08-776-971-4	Sequence 4, Appli

28 72 94.7 19 3 US-09-421-208-30 Sequence 30, Appl
29 72 94.7 20 3 US-09-105-678A-34 Sequence 34, Appl
30 72 94.7 20 3 US-08-776-971-8 Sequence 8, Appl
31 72 94.7 20 3 US-08-776-971-38 Sequence 38, Appl
32 72 94.7 20 3 US-09-421-208-34 Sequence 34, Appl
33 72 94.7 20 4 US-09-560-915-16 Sequence 16, Appl
34 72 94.7 21 3 US-09-105-678A-35 Sequence 35, Appl
35 72 94.7 21 3 US-08-776-971-9 Sequence 9, Appl
36 72 94.7 21 3 US-09-421-208-35 Sequence 35, Appl
37 72 94.7 22 3 US-09-105-678A-36 Sequence 36, Appl
38 72 94.7 22 3 US-08-776-971-10 Sequence 10, Appl
39 72 94.7 22 3 US-08-421-208-36 Sequence 36, Appl
40 72 94.7 25 3 US-08-776-971-111 Sequence 111, App
41 72 94.7 29 3 US-09-105-678A-29 Sequence 29, Appl
42 72 94.7 29 3 US-08-776-971-3 Sequence 3, Appl
43 72 94.7 29 3 US-09-421-208-29 Sequence 29, Appl
44 72 94.7 31 3 US-09-105-678A-7 Sequence 7, Appl
45 72 94.7 31 3 US-09-105-678A-31 Sequence 31, Appl

ALIGNMENTS

RESULT 1
US-09-105-678A-46
; Sequence 46, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-105-678A-46

Query Match 100.0%; Score 76; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 13; Conservative 0; Mismatches 0; Gaps 0;
Qy 1 TPDINPAWASRG 13

us-09-700-643a-2_copy_12_24.ra1

Thu Aug 12 15:25:16 2004

```

Db      1 TPDINPAWYASRG 13

RESULT 2
US-08-776-971-64
; Sequence 64, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
;             Habata, Yugo
;             Kawamata, Yuji
;             Hosoya, Masaki
;             Fujii, Ryo
;             Fukusumi, Shoji
;             Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 64:
US-08-776-971-64
; Query Match 100.0%; Score 76; DB 3; Length 20;
; Best Local Similarity 100.0%; Pred. No. 1.3e-06;
; Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TPDINPAWYASRG 13
Db      1 TPDINPAWYASRG 13

RESULT 3
US-08-776-971-64
; Query Match 100.0%; Score 76; DB 3; Length 20;
; Best Local Similarity 100.0%; Pred. No. 1.3e-06;
; Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TPDINPAWYASRG 13
Db      1 TPDINPAWYASRG 13

RESULT 4
US-09-560-915-18
; Sequence 18, Application US/09560915
; Patent No. 6383764
; GENERAL INFORMATION:
; APPLICANT: Civalli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; TITLE OF INVENTION: Relating to Prolactin Releasing Peptide (PrRP)
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/09/560,915
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo Sapien

```


STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/421,208
FILING DATE: 27-JUN-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 09/105,678
FILING DATE: 26-JUN-1998
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: linear
MOLECULE TYPE: peptide
US-09-421-208-47

Query Match 100.0%; Score 76; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPDINPAWYASRG 13
DB 1 TPDINPAWYASRG 13

RESULT 8
US-09-105-678A-48
Sequence 48, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: linear
MOLECULE TYPE: peptide
US-09-105-678A-48

Query Match 100.0%; Score 76; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPDINPAWYASRG 13
DB 1 TPDINPAWYASRG 13

RESULT 9
US-08-776-971-66
Sequence 66, Application US/08776971B
Patent No. 6228984
GENERAL INFORMATION:
APPLICANT: Hinuma, Shuji
APPLICANT: Habata, Yugo
APPLICANT: Kawamata, Yuji
APPLICANT: Hosoya, Masaki
APPLICANT: Fujii, Ryo
APPLICANT: Fukusumi, Shoji
APPLICANT: Kitada, Chieko
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,971B
FILING DATE: 06-Feb-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03821
FILING DATE: 28-DEC-1996
APPLICATION NUMBER: JP 7/343371
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: JP 8/59419
FILING DATE: 15-MAR-1996
APPLICATION NUMBER: JP 8/211805
FILING DATE: 12-AUG-1996
APPLICATION NUMBER: JP 8/246573
FILING DATE: 18-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 47176
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 56:

SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 66:
US-09-776-971-66

Query Match 100.0%; Score 76; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRG 13
DB 1 TPDINPAWYASRG 13

RESULT 10
US-09-421-208-48
; Sequence 48, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/421,208
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/105,678
FILING DATE: 26-JUN-1998
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-421-208-48

Query Match 100.0%; Score 76; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRG 13
DB 1 TPDINPAWYASRG 13

RESULT 11
US-09-105-678A-9
; Sequence 9, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-9

Query Match 100.0%; Score 76; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRG 13
DB 12 TPDINPAWYASRG 24

RESULT 12
US-09-105-678A-43
; Sequence 43, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA

ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 26-JUN-1998
APPLICATION NUMBER: US/09/105,678A
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-43

Query Match 100.0%; Score 76; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRG 13
Db 12 TPDINPAWYASRG 24

RESULT 13
US-08-776-971-61
Sequence 61, Application US/08776971B
Patent No. 6228984
GENERAL INFORMATION:
APPLICANT: Hinuma, Shuji
Habata, Yugo
Kawamata, Yuji
Hosoya, Masaki
Fujii, Ryo
Fukusumi, Shoji
Kitada, Chieko
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,971B
FILING DATE: 06-Feb-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03821
FILING DATE: 28-DEC-1996
APPLICATION NUMBER: JP 7/343371
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: JP 8/59419

FILING DATE: 15-MAR-1996
APPLICATION NUMBER: JP 8/211805
FILING DATE: 12-AUG-1996
APPLICATION NUMBER: JP 8/246573
FILING DATE: 18-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 47176
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-08-776-971-61

Query Match 100.0%; Score 76; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRG 13
Db 12 TPDINPAWYASRG 24

RESULT 14
US-09-421-208-9
Sequence 9, Application US/09421208
Patent No. 6258561
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/421,208
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/105,678
FILING DATE: 26-JUN-1998
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids

;
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-421-208-9

Query Match 100.0%; Score 76; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRG 13
|||
Db 12 TPDINPAWYASRG 24
|||

RESULT 15
US-09-421-208-43
; Sequence 43, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208

FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-421-208-43

Query Match 100.0%; Score 76; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRG 13
|||
Db 12 TPDINPAWYASRG 24
|||

Search completed: August 12, 2004, 14:52:13
Job time : 7.40698 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:36 ; Search time 5.74419 Seconds
(without alignments)
217.697 Million cell updates/sec

Title: US-09-700-643A-3_COPY_12_24
Perfect score: 79
Sequence: 1 TPDINPAWYTGKG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	83	2 JC7607	prolactin-releasin
2	46	58.2	139	2 H84384	hypothetical prote
3	46	58.2	527	2 T31175	hypothetical prote
4	45	57.0	767	2 T21969	hypothetical prote
5	44.5	56.3	482	1 S40887	RVS167 protein - y
6	44	55.7	398	2 C84780	hypothetical prote
7	44	55.7	449	2 G87789	protein C34G6.7 [i
8	44	55.7	548	2 T47548	hypothetical prote
9	43	54.4	330	2 JE0376	Grb-2 related adap
10	43	54.4	430	1 B69009	conserved hypothet
11	43	54.4	498	2 T09021	beta-glucosidase h
12	43	54.4	517	2 T09022	beta-glucosidase h
13	42	53.2	120	2 T51754	endo-xyloglucan tr
14	42	53.2	269	2 S61555	xyloglucan endo-1,
15	42	53.2	529	2 T51214	related to trichod
16	42	53.2	1425	2 T30811	hepatocyte growth
17	41	51.9	101	2 S30975	gene 30 protein -
18	41	51.9	126	2 B72621	hypothetical prote
19	41	51.9	280	2 S67485	excinnuclease ABC c
20	41	51.9	395	2 T09373	hypothetical prote
21	41	51.9	476	2 G64720	probable amino aci
22	41	51.9	476	2 G90629	probable inner mem
23	41	51.9	476	2 G85480	inner membrane tra
24	41	51.9	925	2 T33732	probable excinnuc
25	41	51.9	935	2 A64608	excinnuclease ABC c
26	41	51.9	939	2 H96961	excinnuclease ABC c
27	41	51.9	940	1 BVECUA	excinnuclease ABC c
28	41	51.9	940	2 H91258	excinnuclease ABC c
29	41	51.9	940	2 A82329	excinnuclease ABC c

30	41	51.9	940	2 D86099	excinnuclease ABC c
31	41	51.9	941	2 F71906	excinnuclease ABC c
32	41	51.9	941	2 B81454	excinnuclease ABC c
33	41	51.9	941	2 AE1017	excinnuclease ABC c
34	41	51.9	943	2 D64057	excinnuclease ABC c
35	41	51.9	943	2 F95021	excinnuclease ABC c
36	41	51.9	943	2 C97893	excinnuclease ABC c
37	41	51.9	945	2 E83117	excinnuclease ABC c
38	41	51.9	947	2 A00040	excinnuclease ABC c
39	41	51.9	948	2 B81883	excinnuclease ABC c
40	41	51.9	948	2 H89848	excinnuclease ABC su
41	41	51.9	949	2 A81138	excinnuclease ABC c
42	41	51.9	952	1 T46550	excinnuclease ABC c
43	41	51.9	953	2 D71645	excinnuclease ABC c
44	41	51.9	955	2 F97861	excinnuclease ABC s
45	41	51.9	956	2 A81761	excinnuclease ABC (

ALIGNMENTS

RESULT 1
JC7607
prolactin-releasing peptide - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C:Accession: JC7607
R:Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Hosoya, T.
Biochem. Biophys. Res. Commun. 281, 53-56, 2001
A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene: Mui
A:Reference number: JC7607; MUID:21092785; PMID:11179599
A:Contents: Spleen
A:Accession: JC7607
A:Molecule type: DNA
A:Residues: 1-83 <YAM>
A:Cross-references: DBJ:AB040612; DBJ:AB040613
C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior pit
release, and stimulation of ACTH secretion from the pituitary.
C:Genetics:
A:Gene: PrRP
A:Introns: 33/1

Query Match 100.0%; Score 79; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TPDINPAWYTGKG 13
DB 33 TPDINPAWYTGKG 45
RESULT 2
H84384
hypothetical protein Vng2339c [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 27-Nov-2001
C:Accession: H84384
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jable
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, R.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: H84384
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-139 <STO>
A:Cross-references: GB:AE004437; NID:gl0581750; PIDN:AAG20444.1; GSPDB:GN00138
C:Genetics:
A:Gene: VNG2339C
C:Superfamily: Archaeoglobus fulgidus hypothetical protein AF0104

Query Match 58.2%; Score 46; DB 2; Length 139;
 Best Local Similarity 61.5%; Pred. No. 3.8;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TPDINPAWYTGGR 13
 |||||
 Db 32 TEDITAAWFTGLG 44

RESULT 3
 T33175
 hypothetical protein C24G6.6 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T33175
 R:Greco, T.; Bradshaw, H.; Keppeler, D.
 submitted to the EMBL Data Library, May 1998
 A:Description: The sequence of C. elegans cosmid C24G6.
 A:Reference number: Z21298
 A:Accession: T33175
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-527 <GRE>
 A:Cross-references: EMBL:AF067936; PIDN:AAC19213.1; GSPDB:GN000023; CESP:C24G6.6
 A:Experimental source: strain Bristol N2; clone C24G6
 C:Genetics:
 A:Gene: CESP:C24G6.6
 A:Map position: 5
 A:Introns: 20/3; 77/1; 129/2; 208/3; 470/2

Query Match 58.2%; Score 46; DB 2; Length 527;
 Best Local Similarity 58.3%; Pred. No. 14;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PDINPAWYTGGR 13
 |||||
 Db 370 PNVLRAWYAGRG 381

RESULT 4
 T21969
 hypothetical protein F38E11.7 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T21969
 R:Matthews, P.
 submitted to the EMBL Data Library, January 1996
 A:Reference number: Z19495
 A:Accession: T21969
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-767 <WIL>
 A:Cross-references: EMBL:Z68342; PIDN:CAA92775.1; GSPDB:GN000022; CESP:F38E11.7
 A:Experimental source: clone F38E11
 C:Genetics:
 A:Gene: CESP:F38E11.7
 A:Map position: 4
 A:Introns: 50/2; 118/1; 139/2; 189/3; 226/1; 248/1; 287/2; 375/2; 432/3; 465/3; 548/1; 6

Query Match 57.0%; Score 45; DB 2; Length 767;
 Best Local Similarity 75.0%; Pred. No. 30;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PDINPAWY 9
 |||||
 Db 747 PDVKPAWY 754

RESULT 5
 S40887
 RV5167 protein - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein YDR388w
 C:Species: Saccharomyces cerevisiae

C:Date: 31-Mar-1992 #sequence_revision 06-Feb-1995 #text_change 21-Jul-2000
 C:Accession: S40887; S69672
 R:Bauer, F.; Urdaci, M.; Aigle, M.; Crouzet, M.
 Mol. Cell. Biol. 13, 5070-5084, 1993
 A:Title: Alteration of a yeast SH3 protein leads to conditional viability with defects in
 A:Reference number: S40887; MUID:93330299; PMID:8336735
 A:Accession: S40887
 A:Molecule type: DNA
 A:Residues: 1-482 <BAU>
 A:Cross-references: EMBL:M92092; NID:g172615; PIDN:AAA35051.1; PID:g172616
 R:Dietrich, F.S.
 submitted to the EMBL Data Library, July 1995
 A:Description: The sequence of S. cerevisiae cosmids 9481, 9509, 9926, 9461, and lambda
 A:Reference number: S69665
 A:Accession: S69672
 A:Molecule type: DNA
 A:Residues: 1-482 <DIE>
 A:Cross-references: EMBL:U32274; NID:g927313; PIDN:AAB64830.1; PID:g927321; GSPDB:GN000004
 C:Genetics:
 A:Gene: SGD:RVS167; MIPS:YDR388w
 A:Cross-references: SGD:S0002796; MIPS:YDR388w
 A:Map position: 4R
 C:Superfamily: RVS167 protein; RVS161 protein homology; SH3 homology
 C:Keywords: transmembrane protein
 F:4-270/Domain: RVS161 protein homology <RVS>
 F:292-422/Region: alanine/glycine/proline-rich
 F:428-477/Domain: SH3 homology <SH3>

Query Match 56.3%; Score 44.5; DB 1; Length 482;
 Best Local Similarity 66.7%; Pred. No. 23;
 Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 1 TPDINPAWYTGGR 12
 |||||
 Db 454 TPDVN-EWWTGR 464

RESULT 6
 C84780
 Hypothetical protein At2g36400 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: C84780
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: C84780
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-398 <STO>
 A:Cross-references: GB:AE002093; NID:g4581140; PIDN:AAD24624.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g36400
 A:Map position: 2

Query Match 55.7%; Score 44; DB 2; Length 398;
 Best Local Similarity 63.6%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 PDINPAWYTGGR 12
 |||||
 Db 130 PHYQPAWYLGGR 140

RESULT 7
 G87789
 protein C34G6.7 [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 C:Accession: G87789

R;anonymus, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998
 A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
 A:Reference number: AV5000; MUID:99069613; PMID:9851916
 A>Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
 A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A:Accession: G87789
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-449 <STO>
 A:Cross-references: GB:chr_I; PIDN:AAB52480.1; PID:gl943842; GSPDB:GN00019; CESP:C34G6.7
 A:Note: contains similarity to an SH3 domain
 C:Genetics:
 A:Gene: C34G6.7
 A:Map position: 1

Query Match 55.7%; Score 44; DB 2; Length 449;
 Best Local Similarity 58.3%; Pred. No. 35;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TPDINPAWYTG 12
 : : : : :
 Db 241 TDESHPHWYTG 252

RESULT 8

T47548
 hypothetical protein F8J2.80 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T47548
 R;Nyakatura, G.; Fartmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichselgartner, M.;
 Mayer, K.F.X.
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: Z24458
 A:Accession: T47548
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-548 <NYA>
 A:Cross-references: EMBL:AL132969
 A:Experimental source: cultivar Columbia; BAC clone F8J2
 C:Genetics:
 A:Map position: 3
 A:Introns: 78/1; 143/1; 242/2; 377/3; 428/2; 447/3; 470/3; 502/3; 532/3
 A>Note: F8J2.80

Query Match 55.7%; Score 44; DB 2; Length 548;
 Best Local Similarity 58.3%; Pred. No. 31;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 PDINPAWYTG 13
 : : : : :
 Db 137 PHQPSPWYWG 148

RESULT 9

JE0376
 Grp-2 related adaptor protein 2 - human
 C:Species: Homo sapiens (man)
 C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 21-Jul-2000
 C:Accession: JE0376
 R;Oiu, M.; Hua, S.; Agrawal, M.; Li, G.; Cai, J.; Chan, E.; Zhou, H.; Luo, Y.; Liu, M.
 Biochem. Biophys. Res. Commun. 253, 443-447, 1998
 A>Title: Molecular cloning and expression of human Grap-2, a novel leukocyte-specific SH
 A:Reference number: JE0376; MUID:99097254; PMID:9878555
 A:Accession: JE0376
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-330 <OIU>
 A:Cross-references: GB:AF102694; NID:G3860192; PIDN:AAD04926.1; PID:G3860193

Query Match 54.4%; Score 43; DB 2; Length 330;
 Best Local Similarity 75.0%; Pred. No. 27;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 5 NPAWYTG 12
 : : : : :
 Db 305 NPSWYTG 312

RESULT 10

B69009
 conserved hypothetical protein MTH1070 - Methanobacterium thermoautotrophicum (strain De
 C:Species: Methanobacterium thermoautotrophicum
 C:Date: 29-Jan-1999 #sequence_revision 29-Jan-1999 #text_change 23-Jul-1999
 C:Accession: B69009
 R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
 ; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, N.
 ; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A>Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
 A:Reference number: A59000; MUID:98037514; PMID:9371463
 A:Accession: B69009
 A>Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-430 <MTH>
 A:Cross-references: GB:AE000877; GB:AE000666; NID:G2622157; PIDN:AAB85559.1; PID:G262216
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH1070
 C:Superfamily: conserved hypothetical protein MTH1070

Query Match 54.4%; Score 43; DB 1; Length 430;
 Best Local Similarity 70.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 DINPAWYTG 12
 : : : : :
 Db 191 DINPEWVAG 200

RESULT 11

T09021
 beta-glucosidase homolog T27E11.60 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 22-Oct-1999
 C:Accession: T09021
 R;Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancr
 submitted to the Protein Sequence Database, June 1999
 A:Reference number: Z16533
 A:Accession: T09021
 A:Molecule type: DNA
 A:Residues: 1-498 <BEV>
 A:Cross-references: EMBL:AL078579; GSPDB:GN00062; ATSP:T27E11.60
 A:Experimental source: cultivar Columbia; BAC clone T27E11
 C:Genetics:
 A:Gene: ATSP:T27E11.60
 A:Map position: 4
 A:Introns: 42/3; 64/1; 80/3; 106/1; 130/1; 159/2; 244/3; 283/2; 355/3; 390/1; 420/2
 C:Superfamily: Agrobacterium beta-glucosidase

Query Match 54.4%; Score 43; DB 2; Length 498;
 Best Local Similarity 54.5%; Pred. No. 40;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPDINPAWYTG 11
 : : : : :
 Db 462 TPKLSASWYTG 472

RESULT 12

T09022
 beta-glucosidase homolog T27E11.70 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 22-Oct-1999
 C:Accession: T09022

R; Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancroft, A.; et al. The Arabidopsis Genome Initiative. Nature 400: 468-472 (1999).
 submitted to the Protein Sequence Database, June 1999
 A: Reference number: Z16533
 A: Accession: T09022
 A: Molecule type: DNA
 A: Residues: 1-517 <BEV>
 A: Cross-references: EMBL:AL078579; GSPDB:GN00062; ATSP:T27E11.70
 A: Experimental source: cultivar Columbia; BAC clone T27E11
 C: Genetics:
 A: Gene: ATSP:T27E11.70
 A: Map position: 4
 A: Introns: 42/3; 64/1; 80/3; 106/1; 130/1; 159/2; 244/3; 283/2; 359/1; 367/3; 397/1; 439/1
 C: Superfamily: Agrobacterium beta-glucosidase

Query Match 54.4%; Score 43; DB 2; Length 517;
 Best Local Similarity 54.5%; Pred. No. 42;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPDINPAWYTG 11
 || : : ||||
 DB 481 TPXLSASWYTG 491

RESULT 13
 T51754
 endo-xyloglucan transferase [imported] - Arabidopsis thaliana (fragment)
 C: Species: Arabidopsis thaliana (mouse-ear cress)
 C: Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 03-Nov-2000
 C: Accession: T51754
 R; Park, J.H.; Oh, S.A.; Kim, Y.H.; Woo, H.R.; Nam, H.G.
 Plant Mol. Biol. 37, 445-454, 1998
 A: Title: Differential expression of senescence-associated mRNAs during leaf senescence in Arabidopsis thaliana
 A: Reference number: Z25447; MUID:98278374; PMID:9617812
 A: Accession: T51754
 A: Status: preliminary; translated from GB/EMBL/DDBJ
 A: Molecule type: mRNA
 A: Residues: 1-120 <PAR>
 A: Cross-references: EMBL:AF035384; PIDN:AAC39467.1
 A: Experimental source: cultivar ecotype Col-O; leaf
 C: Genetics:
 A: Gene: SEN4
 A: Note: induced during leaf senescence
 C: Superfamily: endoxyloglucan transferase
 C: Keywords: phosphotransferase; protein kinase

Query Match 53.2%; Score 42; DB 2; Length 120;
 Best Local Similarity 63.6%; Pred. No. 14;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 DINPAWYTG 13
 || : ||||
 DB 27 DVNVAWNGRG 37

RESULT 14
 S61555
 xyloglucan endo-1,4-beta-D-glucanase (EC 3.2.1.-) precursor - Arabidopsis thaliana
 N; Alternate names: endo-xyloglucan transferase; meri-5 protein; protein F9N11.120
 C: Species: Arabidopsis thaliana (mouse-ear cress)
 C: Date: 23-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 20-Jun-2000
 C: Accession: S61555; S57771; JQ1022; T14087
 R; Kamitani, T.; Tomita, E.; Nishitani, K.
 submitted to the EMBL Data Library, July 1995
 A: Reference number: S61555
 A: Accession: S61555
 A: Molecule type: mRNA
 A: Residues: 1-269 <KAM>
 A: Cross-references: EMBL:D63508; NID:G1805365; PIDN:BAA09783.1; PID:G944810
 A: Note: meri5
 R; Arrowsmith, D.A.; de Silva, J.
 Plant Mol. Biol. 28, 391-403, 1995
 A: Title: Characterisation of two tomato fruit-expressed cDNAs encoding xyloglucan endo-1,4-beta-D-glucanase
 A: Reference number: S57769 MUID:95359399; PMID:7632911

us-09-700-643a-3_copy_12_24.rpr

Thu Aug 12 15:25:20 2004

Db 218 PDFGEWYDG 227

Search completed: August 12, 2004, 14:50:28
Job time : 6.74419 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 3.40116 Seconds
(without alignments)
199.024 Million cell updates/sec

Title: US-09-700-643A-3_COPY_12_24
Perfect score: 79
Sequence: 1 TPDINPAWYTGK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	83	1 PRRP RAT	P81278 rattus norv
2	74	93.7	98	1 PRRP BOVIN	P81264 bos taurus
3	68	86.1	87	1 PRRP HUMAN	P81277 homo sapien
4	44.5	56.3	482	1 R167 YEAST	P39743 saccharomyc
5	43	54.4	322	1 GRP2 MOUSE	O89100 m grb2-rela
6	43	54.4	330	1 GRP2 HUMAN	O75791 h grb2-rela
7	42	53.2	269	1 XT24 ARATH	P24806 a xylogluca
8	41	51.9	101	1 VG30 BPML5	Q05239 mycobacteri
9	41	51.9	476	1 YAAJ ECOLI	P30143 escherichia
10	41	51.9	569	1 UVRB VITST	Q08518 viroescill
11	41	51.9	925	1 UVRB ZYMO	O31151 zymomonas m
12	41	51.9	935	1 UVRB HELPY	P56474 helicobacte
13	41	51.9	940	1 UVRB ECO57	Q8X5U9 escherichia
14	41	51.9	940	1 UVRB ECOL6	Q8F802 escherichia
15	41	51.9	940	1 UVRB ECOLI	P07671 escherichia
16	41	51.9	940	1 UVRB VIBCH	Q9KUW5 vibrio chol
17	41	51.9	940	1 UVRB VIBPA	Q871A0 vibrio para
18	41	51.9	940	1 UVRB VIBVU	Q8CJ33 vibrio vuln
19	41	51.9	941	1 UVRB HELPU	Q921D6 helicobacte
20	41	51.9	941	1 UVRB SALTY	P37434 salmonella
21	41	51.9	942	1 UVRB STRP3	Q8X5Z0 streptococc
22	41	51.9	942	1 UVRB STRP8	Q8N2J2 streptococc
23	41	51.9	942	1 UVRB STRPY	Q99V84 streptococc
24	41	51.9	943	1 UVRB HAEIN	P44410 haemophilus
25	41	51.9	943	1 UVRB PASMU	P57979 pasteurella
26	41	51.9	943	1 UVRB STRMU	P72481 streptococc
27	41	51.9	943	1 UVRB STRPN	Q97AX7 streptococc
28	41	51.9	944	1 UVRB PSEPK	Q88GK7 pseudomonas
29	41	51.9	944	1 UVRB STAEF	Q8CPY9 staphylococ
30	41	51.9	945	1 UVRB PSEAE	Q9BWC0 pseudomonas
31	41	51.9	947	1 UVRB YERPE	Q8Z107 yersinia pe
32	41	51.9	948	1 UVRB NEIMA	Q9JUS4 neisseria m
33	41	51.9	948	1 UVRB STAAW	Q99V16 staphylococ

34	41	51.9	948	1 UVRB STAAW	Q8NX19 staphylococ
35	41	51.9	949	1 UVRB NEIMB	Q9JZ11 neisseria m
36	41	51.9	950	1 UVRB NEIGO	Q50968 neisseria g
37	41	51.9	951	1 UVRB LACPL	Q88Y17 lactobacill
38	41	51.9	952	1 UVRB THETH	Q56242 thermus the
39	41	51.9	953	1 UVRB RICPR	Q92CC3 rickettsia
40	41	51.9	955	1 UVRB RICCN	Q92G31 rickettsia
41	41	51.9	956	1 UVRB LISIN	Q928A5 listeria in
42	41	51.9	956	1 UVRB LISMO	Q8Y4F6 listeria mo
43	41	51.9	957	1 UVRB BACHD	Q9K6V0 bacillus ha
44	41	51.9	957	1 UVRB BACSU	Q34863 bacillus su
45	41	51.9	958	1 UVRB OCEIH	Q8ENJ6 oceanobacil

ALIGNMENTS

RESULT 1					
PRRP RAT					
ID	PRRP RAT	STANDARD;	PRT;	83 AA.	
AC	P81278; O8K3Y0;				
DT	30-MAY-2000 (Rel. 39, Created)				
DT	30-MAY-2000 (Rel. 39, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-releasing peptide PrRP20].				
GN	PRH.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).				
RC	TISSUE=Brain.				
RX	MEDLINE=98268781; Pubmed=9607765;				
RA	Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,				
RA	Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,				
RA	Kurokawa T., Nishimura O., Onda H., Fujino M.,				
RT	"A prolactin-releasing peptide in the brain."				
RL	Nature 393:272-276(1998).				
RN	[2]				
RP	SEQUENCE FROM N.A. (ISOFORM 2).				
RC	STRAIN=Sprague-Dawley; TISSUE=Hypothalamus;				
RA	Anderson S.T., Kokay I.C., Lang T., Grattan D.R., Curlewis J.D.;				
RT	"Quantitation of prolactin-releasing peptide (PrRP) mRNA expression in specific brain regions during the rat oestrous cycle and in lactation."				
RL	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	TISSUE SPECIFICITY.				
RX	MEDLINE=99428652; Pubmed=10498338;				
RA	Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,				
RA	Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,				
RA	Sumino Y., Fujino M.;				
RT	"Tissue distribution of prolactin-releasing peptide (PrRP) and its receptor."				
RL	Regul. Pept. 83:1-10(1999).				
CC	-!- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.				
CC	-!- ALTERNATIVE PRODUCTS:				
CC	Event=Alternative splicing; Named isoforms=2;				
CC	Name1;				
CC	Isoid=P81278-1; Sequence=Displayed;				
CC	Name2;				
CC	Isoid=P81278-2; Sequence=VSP_004370;				
CC	-!- TISSUE SPECIFICITY: Widely expressed, with highest levels in medulla oblongata and hypothalamus.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its				

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR ENBL; AB015418; BAA29026.1; --
 DR ENBL; AF521930; AAM82154.1; --
 DR PIR; JC7607; JC7607. Signal; Cleavage on pair of basic residues;
 KW Hormone; Amidation; Signal; Cleavage on pair of basic residues;
 KW Alternative splicing.
 FT SIGNAL 1 21 BY SIMILARITY.
 FT PEPTIDE 22 52 PROLACTIN-RELEASING PEPTIDE PRP31.
 FT PEPTIDE 33 52 PROLACTIN-RELEASING PEPTIDE PRP20.
 FT PROPEP 57 83
 FT MOD_RES 52 52
 FT VARSPLIC 33 83
 FT TPDPINPAWTGGRG 13
 FT TPDPINPAWTGGRG 45

Query Match 100.0%; Score 79; DB 1; Length 83;
 Best Local Similarity 100.0%; Pred. No. 2.4e-06; Indels 0; Gaps 0;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDPINPAWTGGRG 13
 DB 33 TPDPINPAWTGGRG 45

RESULT 2
 PRP BOVIN STANDARD; PRT; 98 AA.
 AC P81264;
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing
 DE hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-
 DE releasing peptide PrRP20].
 GN PRH.
 OS Bos taurus (Bovine).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Bovinae; Bos.
 CC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.
 RC TISSUE=Brain;
 RX MEDLINE=98268781; PubMed=9607765;
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
 RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
 RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
 RT "A prolactin-releasing peptide in the brain."
 RL Nature 393:272-276(1998).
 CC -!- FUNCTION: Stimulates prolactin (PRL) release and regulates the
 CC expression of prolactin through its receptor GPR10. May stimulate
 CC lactotrophs directly to secrete PRL.
 CC -!- TISSUE SPECIFICITY: Medulla oblongata and hypothalamus.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR ENBL; AB015417; BAA29025.1; --
 KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.
 FT SIGNAL 1 22
 FT PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRP31.
 FT PEPTIDE 33 53 PROLACTIN-RELEASING PEPTIDE PRP20.

FT PROPEP 58 98
 FT MOD_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 98 AA; 10344 MW; 08AC35A13B0FA908 CRC64;
 Query Match 93.7%; Score 74; DB 1; Length 98;
 Best Local Similarity 92.3%; Pred. No. 1.9e-05;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPDPINPAWTGGRG 13
 DB 34 TPDPINPAWTGGRG 46

RESULT 3
 PRP HUMAN STANDARD; PRT; 87 AA.
 AC P81277;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing
 DE hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-
 DE releasing peptide PrRP20].
 GN PRH.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98268781; PubMed=9607765;
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
 RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
 RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
 RT "A prolactin-releasing peptide in the brain."
 RL Nature 393:272-276(1998).
 CC (2)
 RP TISSUE SPECIFICITY.
 RX MEDLINE=99426652; PubMed=10498338;
 RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,
 RA Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,
 RA Sumino Y., Fujino M.;
 RT "Tissue distribution of prolactin-releasing peptide (PrRP) and its
 RT receptor."
 RL Regul. Pept. 83:1-10(1999).
 CC -!- FUNCTION: Stimulates prolactin (PRL) release and regulates the
 CC expression of prolactin through its receptor GPR10. May stimulate
 CC lactotrophs directly to secrete PRL.
 CC -!- TISSUE SPECIFICITY: Medulla oblongata and hypothalamus.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR ENBL; AB015419; BAA29027.1; --
 DR MIN; 502663; --
 DR GO; GO:0005180; F:peptide hormone; TAS.
 KW Hormone; Amidation; Signal.
 FT SIGNAL 1 22 BY SIMILARITY.
 FT PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRP31.
 FT PEPTIDE 34 53 PROLACTIN-RELEASING PEPTIDE PRP20.
 FT PROPEP 58 87
 FT MOD_RES 53 53
 SQ SEQUENCE 87 AA; 9639 MW; 229A2F3F50CF981B CRC64;

Query Match 86.1%; Score 68; DB 1; Length 87;
 Best Local Similarity 84.6%; Pred. No. 0.00016;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

CC site selection mechanism. Binds to actin.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M32092; AAA35051.1; -.
CC EMBL; U32274; AAB64830.1; -.
CC FIR; S40887; S40887.
CC HSSP; F19174; 2HSP.
CC GerMOnline; 140880; -.
CC SGD; S0002796; RVS167.
CC GO; GO:0005857; C:actin cortical patch (sensu Saccharomyces); IDA.
CC GO; GO:0008092; F:cytoskeletal protein binding; IPI.
CC GO; GO:0005897; P:endocytosis; IMP.
CC GO; GO:0007121; P:polar budding; IMP.
CC GO; GO:0006970; P:response to osmotic stress; IMP.
CC InterPro; IPR001452; BAR.
CC InterPro; IPR001452; BAR.dom.
CC Pfam; PF03114; BAR; 1.
CC Pfam; PF00018; SH3; 1.
CC PRINTS; PR00452; SH3DOMAIN.
CC ProDom; PD000066; SH3; 1.
CC SMART; SM00721; BAR; 1.
CC SMART; SM00326; SH3; 1.
CC PROSITE; PS50002; SH3; 1.
CC Cytoskeleton; SH3 domain; Coiled coil; Transmembrane; Actin-binding.
CC DOMAIN 31 64 COILED COIL (POTENTIAL).
CC FT DOMAIN 174 204 COILED COIL (POTENTIAL).
CC FT DOMAIN 292 427 ALA/GLY/PRO-RICH.
CC FT TRANSMEM 344 367 POTENTIAL.
CC FT DOMAIN 421 482 SH3.
CC SEQUENCE 482 AA; 52774 MW; 3FOAB53EBC95A5B CRC64;
CC
CC Query Match 56.3%; Score 44.5; DB 1; Length 482;
CC Best Local Similarity 66.7%; Pred. No. 6.9;
CC Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 1;
CC
CC QY 1 TPDINPAWYTGK 12
CC |||||:|||||
CC 454 TPDVN-EWWTGR 464
CC
CC Db
CC
CC RESULT 5
CC GRP2 MOUSE
CC ID GRP2_MOUSE STANDARD; PRT; 322 AA.
CC AC
CC O89100;
CC DT 30-MAY-2000 (Rel. 39, Created)
CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
CC DE GRB2-related adaptor protein 2 (GADS protein) (Growth factor receptor
CC binding protein) (GRB2) (GRB-2-like protein) (GRB2L) (hematopoietic
CC cell-associated adaptor protein GrpL) (GRB-2-related monocytic adapter
CC protein) (Monocytic adapter) (MONA) (Adapter protein GRID).
CC GN GRAP2 OR MONA OR GADS OR GRB2L OR GRID.
CC OS Mus musculus (Mouse).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC NCBI_TaxID=10090;
CC [1]
CC SEQUENCE FROM N.A.
CC RA Kedra D., Dumaniski J.P.;
CC FT "Cloning of the human and mouse growth factor receptor binding protein
CC like genes."
CC RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC [2]

```

Qy 1 TPDINPAWYTCRG 13
Db 34 TPDINPAWYASRG 46

RESULT 4
R167 YEAST
ID P167 YEAST STANDARD; PRT; 482 AA.
AC P39743;
CT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DT Reduced viability upon starvation protein 167.
GN RVS167 OR YDR388W OR D9509.8.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
ON [1]
RP SEQUENCE FROM N.A.
RC STRAIN=X2180;
RX MEDLINE=93330299; PubMed=8336735;
RX Baur F., Urdaci M., Aigle M., Crouzet M.;
RT "Alteration of a yeast SH3 protein leads to conditional viability
RT with defects in cytoskeletal and budding patterns.";
RL Mol. Cell. Biol. 13:5070-5084 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97313263; PubMed=9169867;
RX Jacq C., Alt-Noerbe J., Andre B., Arnold W., Bahr A., Ballesta J.P.G.,
RX Bargues M., Baron L., Becker A., Biteau N., Bloecker H., Blugnon C.,
RX Boskovic J., Brandt P., Bruckner M., Buitrago M.J., Coster F.,
RX Delaveau T., del Rey F., Dujon B., Eide L.G., Garcia-Cantalejo J.M.,
RX Goffeau A., Gomez-Peris A., Granotier C., Hanemann V., Hankeln I.,
RX Hohelsel J.D., Jaeger W., Jimenez A., Joniaux J.-L., Kraemer C.,
RX Kuester H., Laamanen P., Legros Y., Louis E.J., Moeller-Rieker S.,
RX Monnet A., Moro M., Mueller-Auer S., Nussbaumer B., Paricio N.,
RX Paulin L., Perea J., Perez-Alonso M., Perez-Ortin J.E., Pohl T.M.,
RX Prydz H., Purnelle B., Rasmussen S.W., Remacha M., Revuelta J.L.,
RX Rieger M., Salom D., Saluz H.P., Saiz J.E., Saren A.-M., Schaefer M.,
RX Schafae M., Schmidt E.R., Schneider C., Scholler P., Schwarz S.,
RX Urrestazu L.A., Verhasselt P., Vissers S., Voet M., Volckaert G.,
RX Wagner G., Wambutt R., Wedler E., Wedler H., Woelfel S., Harris D.E.,
RX Bowman S., Brown D., Churcher C.M., Connor R., Dedman K., Gentles S.,
RX Hamlin N., Hunt S., Jones L., McDonald S., Murphy L., Niblett D.,
RX Odell C., Oliver K., Rajandream M.A., Richards C., Shore L.,
RX Walsh S.V., Barrell B.G., Dietrich F.S., Mulligan J.T., Allen E.,
RX Araujo R., Bariles E., Berno A., Carpenter J., Chen E., Cherry J.M.,
RX Chung E., Duncan M., Hunicke-Smith S., Hyman R.W., Komp C.,
RX Lashkari D., Lew H., Lin D., Mosedale D., Nakahara K., Namath A.,
RX Oefner P., Oh C., Petel F.X., Roberts D., Schramm S., Schroeder M.,
RX Shogren T., Shroff N., Winant A., Yelton M.A., Botstein D.,
RX Davis R.W., Johnston M., Andrews S., Brinkman R., Cooper J., Ding H.,
RX Du Z., Favello A., Fulton L., Gatung S., Greco T., Hallsworth K.,
RX Hawkins J., Hillier L.W., Jier M., Johnson D., Johnston L.,
RX Kirsten J., Kucaba T., Langston Y., Latreille P., Le T., Mardis E.,
RX Meneses S., Miller N., Nham M., Pauley A., Peluso D., Rifkin L.,
RX Riles L., Taich A., Trevasakis E., Vignati D., Wilcox L., Wohldman P.,
RX Vaudin M., Wilson R., Waterston R., Albermann K., Hani J., Heumann K.,
RX Kleine K., Mewes H.-W., Zollner A., Zaccaria P.;
RT "the nucleotide sequence of Saccharomyces cerevisiae chromosome IV.";
RN Nature 387:75-78 (1997).
RN [3]
RP ACTIN-BINDING.
RX MEDLINE=95236199; PubMed=7719850;
RX Anberg D.C., Basart E., Botstein D.;
RA "Defining protein interactions with yeast actin in vivo.";
RL Nat. Struct. Biol. 2:28-35 (1995).
CC -!- FUNCTION: Component of a cytoskeletal structure that is required
CC for the formation of endocytic vesicles at the plasma membrane
CC level. Could be implicated in cytoskeletal reorganization in
CC response to environmental stresses and could act in the budding
CC

SEQUENCE FROM N.A.
 RA Bourette R.P., Arnaud S., Myles G.M., Rohrschneider L.R.,
 RA Blanchet J.P., Mouchiroud G.;
 RT "Mona, a novel hematopoietic-specific adaptor interacting with the
 RT macrophage-colony-stimulating factor receptor, is implicated in
 RT monocytic/macrophage development.";
 RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9087328; PubMed=9872323;
 RA Liu S.K., McGlade C.J.;
 RT "Gads is a novel SH2 and SH3 domain-containing adaptor protein that
 RT binds to tyrosine-phosphorylated Shc.";
 RT Oncogene 17:3073-3082(1998).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Law C.-L., Ewings M.K., Chaudhary P.M., Solow S.A., Yun T.J.,
 RA Marshall A.J., Hood L., Clark E.A.;
 RT "GrpL, a Grb2-related adaptor protein, interacts with SLP-76 to
 RT regulate NF-AT activation.";
 RT Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Ellis J.H., Ashman C., Burden M., Kilpatrick K.E., Morse M.A.,
 RA Hamblin P.A.;
 RT "GRID, a novel Grb2-related adaptor protein which interacts with the
 RT activated T cell co-stimulatory receptor CD28.";
 RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA STRAIN=C57BL/6J; TISSUE=Thymus;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: Interacts with SLP-76 to regulate NF-AT activation.
 CC Binds to tyrosine-phosphorylated Shc.
 CC -!- SIMILARITY: Contains 1 SH2 domain.
 CC -!- SIMILARITY: Contains 2 SH3 domains.
 CC -!- SIMILARITY: Belongs to the GRB2 / sem-5 / DRK family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AJ011735; CAA09756.1; -
 DR EMBL; AF055465; AAD08803.1; -
 DR EMBL; AF053405; AAC98669.1; -
 DR EMBL; AF129477; AAD41783.1; -
 DR EMBL; AF236118; AAF60318.1; -
 DR EMBL; BC052496; AAH52496.1; -
 DR HSSP; P29354; IGR1.

DR MGD; MGI:133842; Mona.
 DR InterPro; IPR000980; SH2.
 DR Pfam; PF0001452; SH3.
 DR Pfam; PF00017; SH2; 1.
 DR PRINTS; PF00018; SH3; 2.
 DR PRINTS; PR00401; SH2DOMAIN.
 DR PRINTS; PR00452; SH3DOMAIN.
 DR PRODOM; PD000093; SH2; 1.
 DR PRODOM; PD000066; SH3; 2.
 DR SMART; SM00252; SH2; 1.
 DR SMART; SM00326; SH3; 2.
 DR PROSITE; PS00001; SH2; 1.
 DR PROSITE; PS00002; SH3; 2.
 KW SH2 domain; SH3 domain; Repeat.
 FT DOMAIN 1 56 SH3 1.
 FT DOMAIN 58 149 SH2.
 FT DOMAIN 263 322 SH3 2.
 SQ SEQUENCE 322 AA; 36810 MW; 736311D0640CD3D0 CRC64;
 Query Match 54.4%; Score 43; DB 1; Length 322;
 Best Local Similarity 75.0%; Pred. No. 8.2;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Qy 5 NPAWYTGK 12
 Db 297 NPSWITGR 304
 |||:||||
 |||:||||
 RESULT 6
 GRP2 HUMAN STANDARD; PRT; 330 AA.
 ID O75791; O43726;
 AC 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 18-OCT-2003 (Rel. 42, Last annotation update)
 DE GRB2-related adaptor protein 2 (GADS protein) (Growth factor receptor
 DE binding protein) (GRB2L) (GRB2) (Grb4 adaptor protein) (Grb-40) (GRB-2-like
 DE protein) (GRB2L) (GRB2) (P38) (Hematopoietic cell-associated adaptor
 DE protein GrpL) (Adaptor protein GRID) (SH3-SH2-SH3 adaptor Mona).
 GN GRAP2 OR GADS OR GRB2L OR GRID.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RA Frearson J.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kedra D., Dumanski J.P.;
 RT "Cloning of the human and mouse growth factor receptor binding protein
 RT like genes.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Asada H., Ishi N., Sasaki Y., Endo K., Kasai H., Tanaka N.,
 RA Takeshita T., Tsuchiya S., Konno T., Sugamura K.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99097254; PubMed=9878555;
 RA Qiu M., Hua S., Agrawal M., Li G., Cai J., Chan E., Zhou H., Luo Y.,
 RA Liu M.;
 RT "Molecular cloning and expression of human Grap-2, a novel leukocyte-
 RT specific SH2- and SH3-containing adaptor-like protein that binds to
 RT gab-1.";
 RL Biochem. Biophys. Res. Commun. 253:443-447(1998).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Teramoto T., Nagashima M., Terai S., Thorgeirsson S.S.;
 RT "GrbX, new recruited signaling gene having homology with Grb2.";

RL Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
 RP [6]
 RP SEQUENCE FROM N.A.
 RA Law C.-L., Swings M.K., Chaudhary P.M., Solow S.A., Yun T.J.,
 RA Marshall A.J., Hood L., Clark E.A.;
 RA "GrpL, a Grb2-related adaptor protein, interacts with SLP-76 to
 RT regulate NF-AT activation.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Ellis J.H., Ashman C., Burden M., Kilpatrick K.E., Morse M.A.,
 RA Hamblin P.A.;
 RA "Grid, a novel Grb2-related adapter protein which interacts with the
 RT activated T cell co-stimulatory receptor.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
 RN [8]
 RP SEQUENCE FROM N.A.
 RA Bourette R.P., Arnaud S., Guyot B., Mouchiroud G.;
 RA "Cloning of the human homolog of Mena.";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
 RN [9]
 RP SEQUENCE FROM N.A.
 RA GUYOT B., Arnaud S., Pochirairath P., Bourette R.P., Grasset M.F.,
 RA Rigal D., Mouchiroud G.;
 RA "Genomic organization and restricted expression of the human Mena/Gads
 RT gene suggests regulation by two specific promoters.";
 RL Gene 290:173-179 (2002).
 RN [10]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
 RA Clump M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
 RA Bagghley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
 RA Bird C.P., Blakey S.E., Bridgman A.M., Buck D., Burgess J.,
 RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
 RA Clegg S.M., Corbett V.E., Cole C.G., Collier R.E., Connor R.,
 RA Conroy D., Cobley N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
 RA Evans K.L., Fey J.M., Fleming K., French L., Gartner A.A.,
 RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
 RA Hall R.E., Hall-Tanlyn G., Heathcote R.W., Ho S., Holmes S.,
 RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
 RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
 RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
 RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
 RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
 RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
 RA Soderlund C., Spraggon L., Steward C.A., Sulston J.E., Swann R.M.,
 RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
 RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
 RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
 RA Minoshima S., Kawasaki K., Sakaki T., Asakawa S., Kudoh J.,
 RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuoyama S.,
 RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
 RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
 RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
 RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
 RA Zhan M., Zhang G., Chissoe S., Murray J., Miller N., Minx P.,
 RA Fulton R., Johnson D., Benis G., Bentley D., Bradshaw H., Bourne S.,
 RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
 RA Schoet P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,
 RA Xorf I., Bedell J.A., Hillier L.W., Mardis E., Waterston R.,
 RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
 RA Bucarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
 RA Edgeman L., Kim U.J., Shizuya H., Simon M.I., Dumarski J.P.,
 RA Peyrard M., Kedar D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,
 RA O'Brien K.P., Walkinson P., Bodenteich A., Hartman K., Hu X.,
 RA Khan A.S., Lane L., Tilahun Y., Wright H.;
 RA "The DNA sequence of human chromosome 22.";
 RT

Nature 402:489-495 (1999).
 [11]
 RP SEQUENCE FROM N.A.
 RA TISSUE=lung, and Spleen;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carrinchi P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mollany S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalls D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- FUNCTION: Interacts with SLP-76 to regulate NF-AT activation.
 CC Binds to tyrosine-phosphorylated ehc.
 CC -!- SIMILARITY: Contains 1 SH2 domain.
 CC -!- SIMILARITY: Contains 2 SH3 domains.
 CC -!- SIMILARITY: Belongs to the GRB2 / sem-5 / DRK family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 DR EMBL; Y18051; CAA77021.1; -;
 DR EMBL; AJ011736; CAA09757.1; -;
 DR EMBL; AF042380; AAC69273.1; -;
 DR EMBL; AF102894; AAD04926.1; -;
 DR EMBL; AF090456; AAD13027.1; -;
 DR EMBL; AF129476; AAD1782.1; -;
 DR EMBL; AF236120; AAF60320.1; -;
 DR EMBL; AF121002; AAF31758.1; -;
 DR EMBL; AY069959; AAL58573.1; -;
 DR EMBL; Z82206; CAB05103.1; ALT_SEQ.
 DR EMBL; BC025892; AAK25692.1; -;
 DR EMBL; BC026002; AAK26002.1; -;
 DR PIR; JE0376; J50376.
 DR HSP; P29354; IGRI.
 DR Genew; HGNC:4563; GRAP2.
 DR MIM; 604518; -;
 DR GO; GO:0005737; Cytoplasm; TAS.
 DR GO; GO:0005070; F:SH3/SH2 adaptor protein activity; TAS.
 DR GO; GO:0006960; Pan-microbial humoral response (sensu Inver. . .; TAS.
 DR GO; GO:0007267; P:cell-cell signaling; TAS.
 DR GO; GO:0007265; P:RAS protein signal transduction; TAS.
 DR InterPro; IPR000980; SH2.
 DR Pfam; PF00017; SH2; 1.
 DR Pfam; PF00018; SH2; 2.
 DR PRINTS; PR00401; SH2DOMAIN.
 DR PRINTS; PR00452; SH3DOMAIN.
 DR ProDom; PD000093; SH2; 1.
 DR ProDom; PD000066; SH2; 2.
 DR SMART; SMC0252; SH2; 1.
 DR SMART; SMC0326; SH2; 2.
 DR PROSITE; PS00001; SH2; 1.
 DR PROSITE; PS00002; SH2; 2.
 DR SH2 domain; SH3 domain; Repeat; Polymorphism.
 KW

A.

12
RN SEQUENCE FROM N.A.
RP STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R.; Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.P.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12";
RL Science 277:1433-1474 (1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (potential).
CC -!- SIMILARITY: BELONGS TO THE SODIUM:ALANINE SYMPORTER FAMILY
CC (SAF). STRONG, TO H.INFLUENZAE HI0183.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; D10483; BAB96585.1; -;
DR ENBL; AE000111; AAC73118.1; -;
DR PIR; G64720; G64720.
DR EcoGene; EG1155; Yaaj.
DR InterPro; IPR002293; AA/rel_permease1.
DR InterPro; IPR001463; Na/Ala_symport.
DR Pfam; PF01235; Na/Ala_symport_1.
DR PRINTS; PR00175; NAALASMPORT.
DR TIGRFAMs; TIGR00835; agcs; 1.
DR PROSITE; PS00873; NA_ALANINE_SYMPT; 1.
DR KW Hypothetical protein; Transmembrane; Inner membrane; Transport;
KW Symport; Complete proteome.
FT TRANSVMEM 4 24 POTENTIAL.
FT TRANSVMEM 81 101 POTENTIAL.
FT TRANSVMEM 141 161 POTENTIAL.
FT TRANSVMEM 174 194 POTENTIAL.
FT TRANSVMEM 207 227 POTENTIAL.
FT TRANSVMEM 233 253 POTENTIAL.
FT TRANSVMEM 300 320 POTENTIAL.
FT TRANSVMEM 351 371 POTENTIAL.
FT TRANSVMEM 391 411 POTENTIAL.
FT TRANSVMEM 414 434 POTENTIAL.
SQ SEQUENCE 476 AA; 51662 MW; 2F6EB2E12126E63 CRC64;

Query Match 51.9%; Score 41; DB 1; Length 476;
Best Local Similarity 47.1%; Pred. No. 26;
Matches 8; Conservative 1; Mismatches 2; Indels 6; Gaps 1;

QY 3 DIN-----PAWYTRG 13
Db 120 DVNGQFRGGPAWYMAR 136

RESULT 10
ID UVRA_VITST STANDARD; PRT; 569 AA.
AC Q08518;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE UvrABC system protein A (UvrA protein) (Excinuclease ABC subunit A)
DE (Fragment).
GN UVRA.
OS Vitreoscilla stercoraria.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Vitreoscilla.
OX NCBI_TaxID=61;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;

RX MEDLINE=95151294; PubMed=7765771;
RA Liu S.C., Liu Y.X., Webster D.A., Stark B.C.;
RT "Sequence of the region downstream of the Vitreoscilla hemoglobin
RL gene: vgb is not part of a multigene operon";
RA Appl. Microbiol. Biotechnol. 42:304-308 (1994).
CC -!- FUNCTION: The UvrABC repair system catalyzes the recognition and
CC processing of DNA lesions. UvrA is an ATPase and a DNA-binding
CC protein. A damage recognition complex composed of 2 uvrA and 2
CC uvrB subunits scans DNA for abnormalities. When the presence of a
CC lesion has been verified by uvrB, the uvrA molecules dissociate
CC (by similarity).
CC -!- SUBUNIT: Forms a heterotetramer with uvrB during the search for
CC lesions (by similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -!- SIMILARITY: Belongs to the ABC transporter family. UvrA subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; L21670; AAA75507.1; -;
DR HAMAP; MF 00205; -; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR004602; UvrA.
DR Pfam; PF00005; ABC_tran; 1.
DR PRODOM; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR00630; uvrA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
DR KW SOS response; Excision nuclease; DNA repair; DNA recombination;
KW DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;
KW Zinc-finger.
FT NON_TER 1 1
FT NP_BIND 259 266 ATP (BY SIMILARITY).
FT ZN_FING 359 385 C4-TYPE.
SQ SEQUENCE 569 AA; 61917 MW; 97D7E94D4851BA1F CRC64;

Query Match 51.9%; Score 41; DB 1; Length 569;
Best Local Similarity 72.7%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TPDPNPAWYTG 11
Db 315 TPRSNPATYTG 325

RESULT 11
ID UVRA_ZYMMO STANDARD; PRT; 925 AA.
AC Q311E1;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE UvrABC system protein A (UvrA protein) (Excinuclease ABC subunit A).
GN UVRA.
OS Zymomonas mobilis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Zymomonas.
OX NCBI_TaxID=542;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 31821 / ZM4 / CP4;
RC Kang H.L., Jin S.J., Kang H.S.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: The UvrABC repair system catalyzes the recognition and
CC processing of DNA lesions. UvrA is an ATPase and a DNA-binding
CC protein. A damage recognition complex composed of 2 uvrA and 2

```

CC uvrB subunits scans DNA for abnormalities. When the presence of a
CC lesion has been verified by uvrB, the uvrA molecules dissociate
CC (By similarity).
CC -1- SUBUNIT: Forms a heterotetramer with uvrB during the search for
CC lesions (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the ABC transporter family. UvrA subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF086791; AAC70369.1; -.
CC DR PIR; T33732; T33732.
CC DR HAMAP; MF_00205; atypical; 1.
CC DR InterPro; IPR001439; ABC transporter.
CC DR InterPro; IPR004602; UvrA.
CC DR Pfam; PF00005; ABC_tran; 1.
CC DR ProDom; PD000006; ABC transporter; 1.
CC DR TIGRfams; TIGR00630; uvrA; 1.
CC DR PROSITE; PS00211; ABC TRANSPORTER 1; 2.
CC DR PROSITE; PS00893; ABC TRANSPORTER 2; 1.
CC DR SOS response; Excision nuclease; DNA repair; DNA recombination;
CC DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;
CC Zinc-finger.
CC FT NP_BIND 32 39 ATP (POTENTIAL).
CC FT NP_BIND 644 651 ATP (POTENTIAL).
CC FT ZN_FING 743 769 C4-TYPE.
CC SQ SEQUENCE 925 AA; 101648 MW; 8AF0C36C83712AC5 CRC64;
CC
CC Query Match 51.9%; Score 41; DB 1; Length 925;
CC Best Local Similarity 72.7%; Pred. No. 51;
CC Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
CC
CC QY 1 TPDINPAWYTG 11
CC ||| ||| |||
CC Db 699 TPRSNPATYTG 709
CC
CC RESULT 12
CC UVR_A_HELPY STANDARD; PRT; 935 AA.
CC AC P56474;
CC DT 15-JUL-1998 (Rel. 36, Created)
CC DT 15-JUL-1998 (Rel. 36, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE UvrABC system protein A (UvrA protein) (Excinuclease ABC subunit A).
CC UVR_A OR HP0705.
CC OS Helicobacter pylori (Campylobacter pylori).
CC OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
CC OC Helicobacteraceae; Helicobacter.
CC OX NCBI_TaxID=210;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN=26695 / ATCC 700392;
CC MEDLINE=97394467; PubMed=9252185;
CC RA Tomb J.-P., White O., Kervilange A.R., Clayton R.A., Sutton G.G.,
CC Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
CC Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
CC Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
CC McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
CC Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
CC Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
CC Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
CC Venter J.C.;
CC RA "The complete genome sequence of the gastric pathogen Helicobacter
CC pylori.";
CC RL Nature 388:539-547(1997).
CC -1- FUNCTION: The UvrABC repair system catalyzes the recognition and

```

```

CC processing of DNA lesions. UvrA is an ATPase and a DNA-binding
CC protein. A damage recognition complex composed of 2 uvrA and 2
CC uvrB subunits scans DNA for abnormalities. When the presence of a
CC lesion has been verified by uvrB, the uvrA molecules dissociate
CC (By similarity).
CC -1- SUBUNIT: Forms a heterotetramer with uvrB during the search for
CC lesions (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the ABC transporter family. UvrA subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000583; AAD07755.1; -.
CC DR PIR; A64608; A64608.
CC DR TIGR; HP0705; -.
CC DR HAMAP; MF_00205; -; 1.
CC DR InterPro; IPR003593; AAA ATPase.
CC DR InterPro; IPR003439; ABC transporter.
CC DR InterPro; IPR004602; UvrA.
CC DR Pfam; PF00005; ABC_tran; 1.
CC DR ProDom; PD000006; ABC transporter; 1.
CC DR SMART; SM00382; AAA; 1.
CC DR TIGRfams; TIGR00630; uvrA; 1.
CC DR PROSITE; PS00211; ABC TRANSPORTER 1; 2.
CC DR PROSITE; PS00893; ABC TRANSPORTER 2; 1.
CC DR SOS response; Excision nuclease; DNA repair; DNA recombination;
CC DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;
CC Zinc-finger; Complete proteome.
CC FT NP_BIND 31 38 ATP (POTENTIAL).
CC FT ZN_FING 254 281 C4-TYPE.
CC FT NP_BIND 631 638 ATP (POTENTIAL).
CC FT ZN_FING 731 757 C4-TYPE.
CC SQ SEQUENCE 935 AA; 104060 MW; 6ACF1A675999E141 CRC64;
CC
CC Query Match 51.9%; Score 41; DB 1; Length 935;
CC Best Local Similarity 72.7%; Pred. No. 51;
CC Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
CC
CC QY 1 TPDINPAWYTG 11
CC ||| ||| |||
CC Db 687 TPRSNPATYTG 697
CC
CC RESULT 13
CC UVR_A_ECO57 STANDARD; PRT; 940 AA.
CC AC Q8X5U9;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE UvrABC system protein A (UvrA protein) (Excinuclease ABC subunit A).
CC UVR_A OR Z5657 OR ECS5040.
CC OS Escherichia coli O157:H7.
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC OC Enterobacteriaceae; Escherichia.
CC OX NCBI_TaxID=83334;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN=O157:H7 / EDL933 / ATCC 700927;
CC MEDLINE=21074935; PubMed=11206551;
CC RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
CC Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
CC Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
CC Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
CC Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
CC Welch R.A., Blattner F.R.;
CC RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
CC RT

```

```

RN Nature 409:529-533 (2001).
RP [2]
RC SEQUENCE FROM N.A.
RX STRAIN=O157:H7 / RMD 0509952;
RA MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RL O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -!- FUNCTION: The UvrABC repair system catalyzes the recognition and
CC processing of DNA lesions. UvrA is an ATPase and a DNA-binding
CC protein. A damage recognition complex composed of 2 uvrA and 2
CC uvrB subunits scans DNA for abnormalities. When the presence of a
CC lesion has been verified by uvrB, the uvrA molecules dissociate
CC (By similarity).
CC -!- SUBUNIT: Forms a heterotrimer with uvrB during the search for
CC lesions (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the ABC transporter family. UvrA subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE005638; AAG59256.1; -
CC ENBL: AP002568; BAB38463.1; -
CC PIR: H91258; H91258.
CC HAMAP: MF_00205; - 1.
CC InterPro: IPR003439; ABC transporter.
CC InterPro: IPR004602; UvrA.
CC Pfam: PF00005; ABC tran; 1.
CC ProDom: PD000006; ABC transporter; 1.
CC TIGRfams: TIGR00630; UvrA; 1.
CC PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
CC PROSITE: PS00893; ABC_TRANSPORTER_2; 1.
CC SOS response; Excision nuclease; DNA repair; DNA recombination;
CC DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;
CC Zinc-finger; Complete proteome.
CC NP_BIND 31 38 ATP (POTENTIAL).
CC ZN_FING 253 280 C4-TYPE.
CC ZN_FING 640 647 ATP (POTENTIAL).
CC ZN_FING 740 766 C4-TYPE.
CC SEQUENCE 940 AA; 103884 MW; A20C90C935A0ACEB CRC64;
Query Match 51.9%; Score 41; DB 1; Length 940;
Best Local Similarity 72.7%; Pred. No. 51;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TPDINPAWYTG 11
DB 696 TPRSNDPATYTG 706
RESULT 14
ID UVR_A_ECOL6 STANDARD; PRT; 940 AA.
AC Q8FB02;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE UvrABC system protein A (UvrA protein) (Excinuclease ABC subunit A).
GN UVR_A OR C5048.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;

```

```

RN SEQUENCE FROM N.A.
RP STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RL of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -!- FUNCTION: The UvrABC repair system catalyzes the recognition and
CC processing of DNA lesions. UvrA is an ATPase and a DNA-binding
CC protein. A damage recognition complex composed of 2 uvrA and 2
CC uvrB subunits scans DNA for abnormalities. When the presence of a
CC lesion has been verified by uvrB, the uvrA molecules dissociate
CC (By similarity).
CC -!- SUBUNIT: Forms a heterotrimer with uvrB during the search for
CC lesions (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the ABC transporter family. UvrA subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE016770; AAN83474.1; -
CC HAMAP: MF_00205; - 1.
CC InterPro: IPR003439; ABC transporter.
CC InterPro: IPR004602; UvrA.
CC Pfam: PF00005; ABC tran; 1.
CC ProDom: PD000006; ABC transporter; 1.
CC TIGRfams: TIGR00630; UvrA; 1.
CC PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
CC PROSITE: PS00893; ABC_TRANSPORTER_2; 1.
CC SOS response; Excision nuclease; DNA repair; DNA recombination;
CC DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;
CC Zinc-finger; Complete proteome.
CC NP_BIND 31 38 ATP (POTENTIAL).
CC ZN_FING 253 280 C4-TYPE.
CC ZN_FING 640 647 ATP (POTENTIAL).
CC ZN_FING 740 766 C4-TYPE.
CC SEQUENCE 940 AA; 103882 MW; A20C90C93816ACEB CRC64;
Query Match 51.9%; Score 41; DB 1; Length 940;
Best Local Similarity 72.7%; Pred. No. 51;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TPDINPAWYTG 11
DB 696 TPRSNDPATYTG 706
RESULT 15
ID UVR_A_ECOLI STANDARD; PRT; 940 AA.
AC P07671; P76788;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE UvrABC system protein A (UvrA protein) (Excinuclease ABC subunit A).
GN UVR_A OR D1NE OR B4058 OR SF4146 OR S3583.
OS Escherichia coli, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 623;
RN [1]
RP SEQUENCE FROM N.A.

```

RC SPECIES=E.coli;
RX MEDLINE=86168204; PubMed=3007478;
RA Hsuan I., van Houten B., Thomas D.C., Sancar A.;
RT "Sequences of Escherichia coli uvrA gene and protein reveal two
RL potential ATP binding sites.";
RN J. Biol. Chem. 261:4895-4901(1986).
[2]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=94089392; PubMed=8265357;
RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
RT Daniels D.L.;
RL "Analysis of the Escherichia coli genome. IV. DNA sequence of the
RN region from 89.2 to 92.8 minutes.";
RP Nucleic Acids Res. 21:5408-5417(1993).
[3]
RC SEQUENCE OF 1-25 FROM N.A.
RC SPECIES=E.coli;
RX MEDLINE=83299251; PubMed=6310514;
RA Backendorf C., Brandas J.A., Kartasova T., van de Putte P.;
RT "In vivo regulation of the uvrA gene: role of the '-10' and '-35'
RL promoter regions.";
RN Nucleic Acids Res. 11:5795-5810(1983).
[4]
RC SEQUENCE OF 1-14 FROM N.A.
RC SPECIES=E.coli;
RX MEDLINE=82220077; PubMed=6283374;
RA Sancar A., Sancar G.B., Rupp W.D., Little J.W., Mount D.W.;
RT "LexA protein inhibits transcription of the E. coli uvrA gene in
RL vitro.";
RN Nature 298:96-98(1982).
[5]
RC SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=22272405; PubMed=12384530;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RL Yang J., Yang F., Zhang J., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RN Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RL Yu J.;
RN "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
[6]
RC SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RL Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RN Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RL flexneri serotype 2a strain 2457T.";
RN Infect. Immun. 71:2775-2786(2003).
[7]
RC CHARACTERIZATION.
RC SPECIES=E.coli;
RX MEDLINE=91208117; PubMed=1826851;
RA Myles G.M., Sancar A.;
RT "Isolation and characterization of functional domains of UvrA.";
RN Biochemistry 30:3834-3840(1991).
[8]
RC MUTAGENESIS OF CYS-253.
RC SPECIES=E.coli;
RX MEDLINE=89380205; PubMed=2550431;
RA Navaratnam S., Myles G.M., Strange R.W., Sancar A.;
RT "Evidence from extended X-ray absorption fine structure and site-
RN specific mutagenesis for zinc fingers in UvrA protein of Escherichia
RL coli.";
RN J. Biol. Chem. 264:16067-16071(1989).
CC -!- FUNCTION: The UvrABC repair system catalyzes the recognition and
CC processing of DNA lesions. UvrA is an ATPase and a DNA-binding
CC protein. A damage recognition complex composed of 2 uvrA and 2

CC uvrB subunits scans DNA for abnormalities. When the presence of a
CC lesion has been verified by uvrB, the uvrA molecules dissociate
CC (By similarity).
CC -!- SUBUNIT: Forms a heterotetramer with uvrB during the search for
CC lesions (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- MISCELLANEOUS: Binds about 2 zinc atoms/molecule.
CC -!- SIMILARITY: Belongs to the ABC transporter family. UvrA subfamily.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M13495; AAA24754.1; -.
CC EMBL; U00006; AAC43152.1; -.
CC EMBL; AE000479; AAC77028.1; -.
CC EMBL; X01621; CAA25764.1; -.
CC EMBL; J01721; AAA24753.1; -.
CC EMBL; AE015423; AAN45568.1; -.
CC EMBL; AE016989; AAP18629.1; -.
CC PIR; A23869; BVECUA.
CC ECO2DBASE; H124.0; 6TH EDITION.
CC EcoGene; EG11061; uvrA.
CC HAMAP; MF_00205; -; 1.
CC InterPro; IPR003439; ABC_transporter.
CC InterPro; IPR004602; UvrA.
CC Pfam; PF00005; ABC_tran; 1.
CC ProDom; PD00006; ABC_transporter; 1.
CC TIGRFAMs; TIGR00630; UvrA; 1.
CC PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
CC PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
CC SOS response; Excision nuclease; DNA repair; DNA recombination;
CC DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;
CC Zinc-finger; Complete proteome.
CC NP_BIND 31 38 ATP.
CC ZN_FING 253 280 C4-TYPE.
CC NP_BIND 540 647 ATP.
CC ZN_FING 740 766 C4-TYPE.
CC MUTAGEN 253 253 C->A,H,S: REDUCED ACTIVITY.
CC SEQUENCE 940 AA; 103867 MW; D61AEB6514B860C CRC64;
SQ
Query Match 51.9%; Score 41; DB 1; Length 940;
Best Local Similarity 72.7%; Pred. No. 51;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TPDINPAWYTG 11
DB 696 TPRGNPATYTG 706
Search completed: August 12, 2004, 14:44:53
Job time : 3.40116 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 17.9128 Seconds
(without alignments)
228.984 Million cell updates/sec

Title: US-09-700-643A-3_COPY_12_24

Perfect score: 79

Sequence: 1 TPDINPAWYTGK 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74	93.7	98	Q8WN12	Q8WN12 ovnis aries
2	55	69.6	117	Q9W624	Q9W624 carassius a
3	49	62.0	54	Q726Y1	Q726Y1 homo sapien
4	49	62.0	465	Q60687	Q60687 homo sapien
5	49	62.0	465	Q8W85	Q8W85 homo sapien
6	46	58.2	139	Q9HMX7	Q9HMX7 halobacteri
7	46	58.2	527	Q76383	Q76383 caenorhabdi
8	45	57.0	179	Q20170	Q20170 caenorhabdi
9	45	57.0	419	Q9V276	Q9V276 homo sapien
10	45	57.0	419	Q722V7	Q722V7 homo sapien
11	45	57.0	420	Q72V60	Q72V60 brachydanio
12	44	55.7	196	Q84WLO	Q84WLO arabidopsis
13	44	55.7	380	Q8L8A7	Q8L8A7 arabidopsis
14	44	55.7	397	Q817I2	Q817I2 caenorhabdi
15	44	55.7	398	Q9SJRS	Q9SJRS arabidopsis
16	44	55.7	398	Q8GYS6	Q8GYS6 arabidopsis

17	44	55.7	457	5	O01498	O01498 caenorhabdi
18	44	55.7	467	11	Q8R054	Q8R054 mus musculu
19	44	55.7	468	11	Q8KIF8	Q8KIF8 mus musculu
20	44	55.7	548	10	Q9LFA0	Q9LFA0 arabidopsis
21	44	55.7	1845	17	Q8TTS7	Q8TTS7 methanocarc
22	43	54.4	86	16	Q7U4H3	Q7U4H3 synchococc
23	43	54.4	152	8	Q8WKZ7	Q8WKZ7 c chloropia
24	43	54.4	152	10	Q39562	Q39562 chlamydomon
25	43	54.4	250	14	Q8BFE9	Q8BFE9 uncultured
26	43	54.4	330	4	Q9NR37	Q9NR37 homo sapien
27	43	54.4	430	17	O27142	O27142 methanobact
28	43	54.4	498	10	Q9STP4	Q9STP4 arabidopsis
29	43	54.4	501	13	Q8JI36	Q8JI36 brachydanio
30	43	54.4	508	10	Q93Z14	Q93Z14 arabidopsis
31	43	54.4	517	10	Q9STP3	Q9STP3 arabidopsis
32	43	54.4	540	16	Q886T3	Q886T3 pseudomonas
33	43	54.4	630	4	Q9NXS8	Q9NXS8 homo sapien
34	43	54.4	813	17	Q8ZXH2	Q8ZXH2 pyrobaculum
35	43	54.4	1595	4	Q8IVG5	Q8IVG5 homo sapien
36	42.5	53.8	954	16	Q7WP29	Q7WP29 bordetella
37	42.5	53.8	954	16	Q7WIC4	Q7WIC4 bordetella
38	42.5	53.8	954	16	Q7W0E3	Q7W0E3 bordetella
39	42	53.2	222	16	Q7WP11	Q7WP11 bordetella
40	42	53.2	222	16	Q7VUE5	Q7VUE5 bordetella
41	42	53.2	259	16	Q7WIA7	Q7WIA7 bordetella
42	42	53.2	269	10	Q8LDQ0	Q8LDQ0 arabidopsis
43	42	53.2	468	16	Q88WY8	Q88WY8 lactobacill
44	42	53.2	501	16	Q88554	Q88554 pseudomonas
45	42	53.2	505	16	Q88RQ2	Q88RQ2 pseudomonas

ALIGNMENTS

RESULT 1

Q8WN12 PRELIMINARY; PRT; 98 AA.
 ID Q8WN12;
 AC Q8WN12;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Preprolactin-releasing peptide.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Curlewis J.D., Kusters D.H.L., Barclay J.L., Anderson S.T.;
 RT "Prolactin-releasing peptide (PrP) in the ewe: cDNA cloning, mRNA
 RT distribution and effects on prolactin secretion in vitro and in
 RT vivo."
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF450453; AAI47178.1; -;
 SQ SEQUENCE 98 AA; 10513 MW; 2A53331ED62CAAB5 CRC64;

Query Match 93.7%; Score 74; DB 6; Length 98;
 Best Local Similarity 92.3%; Pred. No. 0.00015;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGK 13
 |||||

Db 34 TPDINPAWYAGRG 46
 |||||

RESULT 2

Q9W624 PRELIMINARY; PRT; 117 AA.
 ID Q9W624;
 AC Q9W624;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)


```

SQ SEQUENCE 527 AA; 59805 MW; 9FBB1FB94437C5CB CRC64;
Query Match 58.2%; Score 46; DB 5; Length 527;
Best Local Similarity 58.3%; Pred. No. 36;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PDINPAWYTCRG 13
   :: ||| |||
Db 370 PNVLAWYAGRG 381

RESULT 8
Q20170 PRELIMINARY; PRT; 179 AA.
AC Q20170;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-OCT-2003 (TEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE C. elegans WRT-3 protein (Corresponding sequence F39E11.7).
GN WRT-3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
   [1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
   investigating biology.";
RL Science 282:2012-2018 (1998).
   [2]
RN RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Matthews P.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z68342; CAA92775.2; -.
SQ SEQUENCE 179 AA; 20965 MW; F9FA6891A2BF2BDD CRC64;

Query Match 57.0%; Score 45; DB 5; Length 179;
Best Local Similarity 75.0%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PDINPAWY 9
   :: ||| |||
Db 159 PDVKPAWY 166

RESULT 9
Q9Y276 PRELIMINARY; PRT; 419 AA.
AC Q9Y276;
DT 01-NOV-1999 (TEMBLrel. 12, Created)
DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DE H-BCS1 (BCS1 (Yeast homolog)-like).
GN BCS1 OR BCS1L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
   [1]
RN RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99097350; PubMed=9878253;
RA Petruzzella V., Tiranti V., Fernandez P., Ianna P., Carrozzo R.,
   Zeviani M.;
RT "Identification and characterization of human cDNAs specific to BCS1,
   Prf12, SCO1, COX15, and COX11, five genes involved in the formation
   and function of the mitochondrial respiratory chain.";
RL Genomics 54:494-504 (1998).
   [2]

```

RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96207227; PubMed=8619474;
RA Andersson B., Wenland M.A., Ricafrente J.Y., Liu W., Gibbs R.A.;
RT "A 'double adaptor' method for improved shotgun library
construction.";
RL Anal. Biochem. 236:107-113(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX Yu W., Andersson B., Worley K.C., Muzny D.M., Ding Y., Liu W.,
RA Ricafrente J.Y., Wenland M.A., Lennon G., Gibbs R.A.;
RT "Large-scale concatenation cDNA sequencing.";
RL Genome Res. 7:353-358(1997).
RN [4]
RP SEQUENCE FROM N.A.
RA de Lonlay P., Valnot I., Barrientos A., Gorbatyuk M., Tzagoloff A.,
RC Benayoun E., Chretien D., Kadhom N., Lombes A., Ogier de Baulny H.,
RA Naudet P., Munnich A., Rustin P., Rotig A.;
RT "Mutations in bcs1, a mitochondrial respiratory chain assembly gene,
are responsible for the complex III deficiency of patients with
tubulopathy and liver failure.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=22233229; PubMed=12215968;
RA Visapaa I., Fellman V., Vesa J., Dasvarma A., Hutton J.L., Kumar V.,
RA Payne G.S., Makarov M., Van Coster R., Turnbull D.M.,
RA Suomalainen A., Peltonen L.;
RT "GRACILE syndrome, a lethal metabolic disorder with iron overload, is
caused by a point mutation in BCS1L.";
RL Am. J. Hum. Genet. 71:863-876(2002).
DR EMBL; AF026849; AA008638.1; -;
DR EMBL; AF038195; AA097365.1; -;
DR EMBL; AF346835; AA029417.1; -;
DR EMBL; BC000416; AA000416.1; -;
DR EMBL; BC007500; AA007500.1; -;
DR EMBL; AF516670; AA005490.1; -;
DR Genew; HGNC:1020; BCS1L.
DR GO; GO:0005750; C:respiratory chain complex III (sensu Eukarya); TAS.
DR GO; GO:0015980; P:energy derivation by oxidation of organic c. . .; TAS.
DR GO; GO:0006461; P:protein complex assembly; TAS.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR Pfam; PF000004; AAA; 1.
DR SMART; SM00382; AAA; 1.
KW Hypothetical protein; ATP-binding.
SQ SEQUENCE 419 AA; 47534 MW; 7F0F98BA62F2CBB8 CRC64;
Query Match 57.0%; Score 45; DB 4; Length 419;
Best Local Similarity 77.8%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 5 NPAWYTGRC 13
Db 211 NPKWYTDRC 219
RESULT 10
Q7Z2V7 PRELIMINARY; PRT; 419 AA.
AC Q7Z2V7
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein DKF2666K10234.
GN DKF2666K10234.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RA Ottenwaelder B., Obermaier B., Deutschenbaue S., Mewes H.W., Weil B.,
RA Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX571752; CAB11877.1; -;
KW Hypothetical protein.
SQ SEQUENCE 419 AA; 47564 MW; 7F1BC9FF62F2CBB8 CRC64;
Query Match 57.0%; Score 45; DB 4; Length 419;
Best Local Similarity 77.8%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 5 NPAWYTGRC 13
Db 211 NPKWYTDRC 219
RESULT 11
Q7Z2V6 PRELIMINARY; PRT; 420 AA.
AC Q7Z2V6
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to BCS1-like (Yeast).
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Body;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC045990; AA045990.1; -;
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR Pfam; PF000004; AAA; 1.
DR SMART; SM00382; AAA; 1.
KW SMART; SM00382; AAA; 1.
SQ SEQUENCE 420 AA; 47448 MW; 370C84A1BF398D24 CRC64;
Query Match 57.0%; Score 45; DB 13; Length 420;
Best Local Similarity 77.8%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 5 NPAWYTGRC 13
Db 211 NPKWYTDRC 219
RESULT 12
Q84WLO PRELIMINARY; PRT; 196 AA.
AC Q84WLO
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Arabidopsis thaliana (Mouse-ear cress).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]


```
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Kim J.H., Kende H.;
RT "A novel class of transcription activators interacting with putative
  co-activators in Arabidopsis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006919; AAD24624.1; -
DR EMBL; AY102636; AAM52878.1; -
DR PIR; C84780; C84780.1; -
SQ SEQUENCE 398 AA; 43707 MW; 2A6E15497305AF74 CRC64;

Query Match      55.7%; Score 44; DB 10; Length 398;
Best Local Similarity 63.6%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      2 PDINPWTGR 12
Db      130 PHYQPAWYLG 140

Search completed: August 12, 2004, 14:49:03
Job time : 17.9128 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:35 ; Search time 27.5872 Seconds
(without alignments)
133.146 Million cell updates/sec

Title: US-09-700-643A-3_COPY_12_24

Perfect score: 79

Sequence: 1 TPDINPAWVTGRG 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	20	AAW31387	Aaw31387 Rat type
2	79	100.0	20	AAW95175	Aaw95175 Murine pi
3	79	100.0	20	AAW97234	Aaw97234 Rat type
4	79	100.0	20	AAW49302	Aay49302 19P2 liiga
5	79	100.0	20	AAW10358	Aab10358 Rat oxyto
6	79	100.0	20	AAW90994	Aab90994 Prolactin
7	79	100.0	20	AAW62527	Aag62527 Rat CRH r
8	79	100.0	20	AAW26403	Aae26403 Rat PrRP-
9	79	100.0	20	AAW60840	ABU60840 Peptide p
10	79	100.0	21	AAW31388	Aaw31388 Rat type
11	79	100.0	21	AAW10359	Aab10359 Rat oxyto
12	79	100.0	21	AAW62528	Aag62528 Rat CRH r
13	79	100.0	21	AAW60841	ABU60841 Peptide p
14	79	100.0	22	AAW31389	Aaw31389 Rat type
15	79	100.0	22	AAW10360	Aab10360 Rat oxyto
16	79	100.0	22	AAW62529	Aag62529 Rat CRH r
17	79	100.0	22	AAW60842	ABU60842 Peptide p
18	79	100.0	31	AAW31384	Aaw31384 Rat type
19	79	100.0	31	AAW95174	Aaw95174 Murine pi
20	79	100.0	31	AAW95173	Aaw95173 Murine pi
21	79	100.0	31	AAW87614	Aaw87614 Rat 19P2
22	79	100.0	31	AAW97233	Aaw97233 Rat type
23	79	100.0	31	AAW49232	Aay49232 19P2 liiga
24	79	100.0	31	AAW87504	Aay87504 Rat prola
25	79	100.0	31	AAW10355	Aab10355 Rat oxyto

26	79	100.0	31	AAW90993	Aab90993 Prolactin
27	79	100.0	31	AAW62524	Rat CRH r
28	79	100.0	31	AAW26400	Rat PrRP-
29	79	100.0	31	ABU60826	Peptide p
30	79	100.0	31	ABU60837	Peptide p
31	79	100.0	32	AAW31385	Rat type
32	79	100.0	32	AAW10356	Rat oxyto
33	79	100.0	32	AAW62525	Rat CRH r
34	79	100.0	32	ABU60838	Peptide p
35	79	100.0	33	AAW31386	Rat type
36	79	100.0	33	AAW10357	Rat oxyto
37	79	100.0	33	AAW62526	Rat CRH r
38	79	100.0	33	ABU60839	Peptide p
39	79	100.0	82	AAW95172	Murine pi
40	79	100.0	83	AAW31383	Rat type
41	79	100.0	83	AAW97225	Rat type
42	79	100.0	83	AAW10354	Rat oxyto
43	79	100.0	83	AAW62523	Rat CRH r
44	75	94.9	31	AAW73370	bPrP31 p
45	74	93.7	19	AAW31370	Bovine G

ALIGNMENTS

RESULT 1

AAW31387
ID AAW31387 standard; peptide; 20 AA.

AC AAW31387;

DT 27-AUG-2003 (revised)

DT 06-APR-1998 (first entry)

XX

DE DE Rat type G protein-coupled receptor ligand fragment 4.

XX G protein-coupled receptor; ligand binding; pharmaceutical; modulator;

KW pituitary; central nervous system; pancreas; prophylactic;

KW therapeutic agent.

XX Rattus sp.

OS WO9724436-A2.

XX PD 10-JUL-1997.

XX PF 26-DEC-1996; 96WO-JP003821.

XX PR 28-DEC-1995; 95JP-00343371.

XX PR 15-MAR-1996; 96JP-00059419.

XX PR 12-AUG-1996; 96JP-00211805.

XX PA 18-SEP-1996; 96JP-00246573.

XX (TAKE) TAKEDA CHEM IND LTD.

XX PI Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukushima S;

XX PI Kitada C;

XX DR WPI; 1997-363672/33.

XX DR N-PSDB; AAV02424.

XX PT Ligand peptide for G protein-coupled receptor - acts by modulating

XX PT function in the central nervous system, pancreas and pituitary gland.

XX PS Claim 2; Page 180; 258pp; English.

XX CC This sequence represents a peptide fragment from a novel rat type ligand

XX CC polypeptide corresponding to amino acid residues 33 to 52 of the sequence

XX CC represented in AAW31383 and is used in an assay to monitor ligand binding

XX CC to the G protein-coupled receptor protein. Pharmaceutical compositions

XX CC containing this ligand may be used as a pituitary function modulator, a

XX CC central nervous system modulator or a pancreatic function modulator. This

XX CC ligand could have specific applications as a prophylactic or therapeutic

CC agent for dementia, depression, hyperkinetic syndrome, disturbance of
 CC consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone
 CC secretory disease, hyper- and polyphagia, hyperlipidaemia,
 CC hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,
 CC ankyloblepharal degeneration, acute myocardial infarction, infertility,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis and/or oligogalactia. Assays can also be developed to screen
 CC compounds which are capable of altering the binding activity of the
 CC ligand affecting activation of the G protein-coupled receptor protein.
 CC (Updated on 27-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 20 AA;

Query Match 100.0%; Score 79; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.3e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGGR 13
 |||||
 Db 1 TPDINPAWYTGGR 13

RESULT 2
 AAW95175
 ID AAW95175 standard; protein; 20 AA.

XX AC AAW95175;

DT 10-MAR-1999 (first entry)

DE Murine pituitary-derived ligand polypeptide antigenic epitope.

XX Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
 KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
 KW Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
 KW gene therapy; transgenic animal; epitope.

XX Mus sp.

XX WO9849295-A1.

XX 05-NOV-1998.

XX 27-APR-1998; 98WO-JP001923.

XX 28-APR-1997; 97JP-00109974.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Hinuma S, Fukusumi S;

XX WPI; 1999-009423/01.

PT New polypeptide ligand for orphan G protein coupled receptors - used for
 PT treating disorders of central nervous system, pituitary and pancreas, and
 PT for drug screening.

XX Disclosure; Page 26; 206pp; English.

CC The invention relates to a murine pituitary-derived ligand polypeptide
 CC which is a ligand for the G-protein coupled orphan receptor designated
 CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing
 CC the ligand polypeptide encoding DNA are used to produce a recombinant
 CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate
 CC function of the pituitary, central nervous system, pancreas and other
 CC tissues and can be used to screen for agents that modulate binding of the
 CC polypeptide to the receptor; to quantify the amount of receptor in a
 CC sample and to raise antibodies. They may also be used therapeutically.

CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's
 CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;
 CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;
 CC rheumatoid arthritis; epilepsy and many others, also to improve post-
 CC operative nutritional status and as vasopressor. Transgenic animals
 CC carrying the ligand polypeptide encoding DNA or its mutagen are used to
 CC study the function of the polypeptide-expressing genes, as models of
 CC disease, for drug screening and as source of cell lines. The ligand
 CC polypeptide DNA is used as a source of probes and primers to identify
 CC related sequences; in receptor-binding assays; for production of Ab and
 CC antisera; in drug development; for gene therapy and to develop transgenic
 CC animals. Sequences AAW95174 to AAW95178 represent antigenic epitopes
 CC which can be used for the preparation of anti-ligand polypeptide antibody
 XX
 SQ Sequence 20 AA;

Query Match 100.0%; Score 79; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.3e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGGR 13
 |||||
 Db 1 TPDINPAWYTGGR 13

RESULT 3
 AAW97234
 ID AAW97234 standard; peptide; 20 AA.

XX AC AAW97234;

DT 06-MAY-1999 (first entry)

DE Rat type ligand polypeptide fragment.

XX Rat type ligand; modulation; prolactin secretion;
 KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst carcinogenesis;
 KW menopausal syndrome; euthyroid; hypometabolism; lactation;
 KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;
 KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;
 KW acromegaly; Chiari-Frömmel syndrome; Argonz-del Castillo syndrome;
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;
 KW contraceptive; placental function; choriocarcinoma; hydattid mole;
 KW irruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;
 KW abnormal lipidmetabolism; oxytocia.

XX Rattus sp.

XX WO9858962-A1.

XX 30-DEC-1998.

XX 22-JUN-1998; 98WO-JP002765.

XX 23-JUN-1997; 97JP-00165437.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Hinuma S, Kawamata Y, Fujii R, Matsumoto H;

XX WPI; 1999-105614/09.

PT Use of G protein-coupled receptor ligands - for modulating prolactin
 PT secretion or placental function, e.g. for treating menopausal syndrome,
 PT tumours, autoimmune disease or abnormal pregnancy.

XX Claim 3; Page 154; 241pp; English.

CC The present sequence represents a rat type ligand fragment. It is used in
 CC the course of the invention. The specification describes an agent for
 CC modulating prolactin secretion which comprises a ligand polypeptide or a
 CC salt, for a G protein-coupled receptor (GPCR) protein. The agents for
 CC promoting prolactin secretion can be used for treating or preventing

CC hypocoarixianism, gonocyst cacogenesis, menopausal syndrome, euthyroid or
 CC hypermetabolism. They can be used for promoting lactation in a domestic
 CC mammal and as an aphrodisiac. The agents for inhibiting prolactin
 CC secretion can be used for treating or preventing pituitary adenomatosis,
 CC brain tumour, emmenopathy, autoimmune disease, prolactinoma,
 CC infertility, impotence, amenorrhea, galactorrhea, acromegaly, Chiari-
 CC Frommel syndrome, Argonz-del Castillo syndrome, Forbes-Albright syndrome,
 CC lymphoma, Sheehan syndrome or dyszoospemia. The inhibitory agents can
 CC also be used as contraceptives. The agents for modulating placental
 CC function can be used for treating or preventing choriocarcinoma, hydatid
 CC mole, irruption mole, abortion, unthrifty fetus, abnormal
 CC saccharometabolism, abnormal lipidmetabolism or oxytocia
 CC
 XX Sequence 20 AA;

Query Match 100.0%; Score 79; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.3e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGK 13
 |||||
 Db 1 TPDINPAWYTGK 13

RESULT 4
 AAY49302
 ID AAY49302 standard; peptide; 20 AA.
 XX
 AC AAY49302;

DT 22-FEB-2000 (first entry)
 XX
 DE 19P2 ligand peptide fragment.
 XX
 KW Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;
 KW pituitary; regulatory mechanism; central nervous system; pancreatic.
 XX
 OS Rattus sp.

XX Key Location/Qualifiers
 FH Modified-site 20
 FT /note="C-terminal amide"
 XX
 XX WO9960112-A1.
 XX
 XX 25-NOV-1999.

XX 20-MAY-1999; 99WO-JP002650.
 XX
 XX 21-MAY-1998; 98JP-00140293.
 XX
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX
 XX Matsumoto H, Kitada C, Hinuma S;
 XX
 XX WPI; 2000-039381/03.

XX New monoclonal antibodies, useful in diagnosis, as drugs and in studying
 XX diseases related to ligand abnormality.
 XX
 XX Disclosure; Page 27; 73pp; Japanese.

XX The invention provides a monoclonal antibody which has a specific
 XX reaction with the part peptide of the C-terminal of 19P2 ligand or its
 XX derivative. The antibodies can be used in diagnosis or to treat or
 XX prevent diseases associated with abnormality in the pituitary function
 XX regulatory mechanism (e.g. promotion of prolactin secretion), central
 XX nervous regulatory mechanism, and pancreatic function regulatory
 XX mechanism. The antibody-based immunoassay can also be applied in
 XX clarifying the physiological functions of the ligand and its derivative.
 XX Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand
 XX
 XX Sequence 20 AA;

Query Match 100.0%; Score 79; DB 3; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.3e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGK 13
 |||||
 Db 1 TPDINPAWYTGK 13

RESULT 5
 AAB10358
 ID AAB10358 standard; peptide; 20 AA.

XX
 AC AAB10358;

DT 24-NOV-2000 (first entry)
 XX
 DE Rat oxytocin secretion promoting peptide SEQ ID NO: 21.

XX Rat; oxytocin secretion promoter; G protein-coupled receptor protein;
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;
 KW veterinary medicine; milk production.
 XX
 OS Rattus sp.

XX WO200038704-A1.
 XX
 XX 06-JUL-2000.
 XX
 XX 22-DEC-1999; 99WO-JP007199.
 XX
 XX 25-DEC-1998; 98JP-00369585.
 XX
 XX (TAKE) TAKEDA CHEM IND LTD.

XX Matsumoto H, Kitada C, Hinuma S;
 XX
 XX WPI; 2000-452298/39.
 XX
 XX Physiologically-active polypeptide recognized as ligand by G protein-
 XX coupled receptor protein, for promoting secretion of oxytocin, as drugs
 XX for diseases relating to oxytocin secretion and in veterinary medicine.
 XX
 XX Claim 5; Page 58; 72pp; Japanese.

XX This invention describes a novel oxytocin secretion-regulating agent
 XX which contains a ligand peptide or its salt for the G protein-coupled
 XX receptor protein. It is useful in the form of drugs for ameliorating,
 XX preventing and treating diseases relating to oxytocin secretion e.g. weak
 XX pains and atonic bleeding, before and after expulsion of placenta,
 XX uterine recovery failure, caesarean section, stoppage of artificial
 XX fertilization or galactostasis and is also applicable in veterinary
 XX medicine for promoting milk production in cow, goat and pig. This
 XX sequence represents a rat peptide which acts as an oxytocin secretion
 XX promoter

XX Sequence 20 AA;

Query Match 100.0%; Score 79; DB 3; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.3e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGK 13
 |||||
 Db 1 TPDINPAWYTGK 13

RESULT 6
 AAB90994
 ID AAB90994 standard; peptide; 20 AA.

XX

AC AAB90994;
 XX
 XX DT 22-JUN-2001 (first entry)
 XX
 DE Prolactin releasing peptide SEQ ID NO:169.
 XX
 XX KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
 XX KW blood component; modification; succinimidy; maleimido group; amino;
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
 XX
 OS Homo sapiens.
 OS Synthetic.
 OS
 PN WO200069900-A2.
 XX
 XX PD 23-NOV-2000.
 XX
 XX PF 17-MAY-2000; 2000WO-US013576.
 XX
 XX PR 17-MAY-1999; 99US-0134406P.
 PR 10-SEP-1999; 99US-0153406P.
 PR 15-OCT-1999; 99US-0159783P.
 XX
 XX PA (CONJ-) CONJUCHEM INC.
 XX
 XX PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
 XX WPI; 2001-112059/12.
 XX
 XX PT Modifying and attaching therapeutic peptides to albumin prevents
 PT peptidase degradation, useful for increasing length of in vivo activity.
 PS Disclosure; Page 245; 73pp; English.
 XX
 XX The present invention describes a modified therapeutic peptide (I)
 CC comprising a therapeutically active amino acid region (III) and a
 CC reactive group (II) (e.g. succinimidy and maleimido groups) attached to
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
 CC factors and neurotransmitters, to protect them from peptidase activity in
 CC vivo for the treatment of various disorders. Endogenous therapeutic
 CC peptides are not suitable as drug candidates as they require frequent
 CC administration due to rapid degradation by peptidases in the body.
 CC Modifying and attaching therapeutic peptides to albumin prevents or
 CC reduces the action of peptidases to increase length of activity (half
 CC life) and specificity as bonding to large molecules decreases
 CC intracellular uptake and interference with physiological processes.
 CC AAB90829 to AAB92441 represent peptides which can be used in the
 CC exemplification of the present invention
 XX
 XX SQ Sequence 20 AA;
 Query Match 100.0%; Score 79; DB 4; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.3e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TPDINPAWYTGK 13
 DB 1 TPDINPAWYTGK 13
 RESULT 7
 AAG62527
 ID AAG62527 standard; peptide; 20 AA.
 XX
 AC AAG62527;
 XX
 XX DT 24-AUG-2001 (first entry)
 XX
 XX DE Rat CRH releasing protein related peptide SEQ ID NO: 21.
 XX
 KW Rat; corticotrophin releasing hormone; CRH; G protein receptor ligand;
 KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;
 KW Addison's disease; adrenal gland hyperfunction; obesity.
 XX
 OS Rattus sp.
 OS
 PN WO200135984-A1.
 XX
 XX PD 25-MAY-2001.
 XX
 XX PF 17-NOV-2000; 2000WO-JP008119.
 XX
 XX PR 18-NOV-1999; 99JP-00327900.
 PR 26-SEP-2000; 2000JP-00297073.
 XX
 XX PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 XX PI Kitada C, Matsumoto H, Hinuma S;
 XX WPI; 2001-355552/37.
 XX
 XX PT Use of G protein receptor ligand or peptide for controlling corticotropin
 PT releasing hormone secretion.
 XX
 XX PS Claim 4; Page 70; 90pp; Japanese.
 XX
 XX The present sequence describes a method of controlling the secretion of
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein
 CC receptor ligand. This can be used to control the secretion of CRH and is
 CC useful as an analgesic or for treating, preventing or ameliorating
 CC diseases associated with CRH secretion such as hyperaldosteronism,
 CC hypercortisolaemia, secondary or chronic hypoadrenocorticism, Addison's
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The
 CC present sequence is a peptide used in the exemplification of the
 CC invention
 XX
 XX SQ Sequence 20 AA;
 Query Match 100.0%; Score 79; DB 4; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.3e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TPDINPAWYTGK 13
 DB 1 TPDINPAWYTGK 13
 RESULT 8
 AAE26403
 ID AAE26403 standard; peptide; 20 AA.
 XX
 XX AAE26403;
 XX
 XX DT 13-DEC-2002 (first entry)
 XX
 XX DE Rat PrRP-31 C-terminal peptide, PrRP-20.
 XX
 XX KW Rat; wakefulness; sleep disorder; prolactin releasing peptide receptor;
 KW PrRP; GPR10; therapy; epilepsy; narcolepsy; sleepiness; sleep apnoea;
 KW insomnia; idiopathic hypersomnia; psychogenic hypersomnia; seizure;
 KW anticonvulsant.
 XX
 OS Rattus sp.
 OS
 PN US2002037533-A1.
 XX
 XX PD 28-MAR-2002.
 XX
 XX PF 17-AUG-2001; 2001US-00932161.
 XX
 XX PR 28-APR-2000; 2000US-00560915.
 XX


```

Query Match          100.0%;   Score 79;   DB 6;   Length 20;
Best Local Similarity 100.0%;   Pred. No. 4 3e-06;
Matches 13;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy      1  TPDINPAWYTCRG 13
        | | | | | | | | | |
Db      1  TPDINPAWYTCRG 13

RESULT 10
AAW31388
ID AAW31388 standard; peptide: 21 AA.

```

XX
XX
DT 27-AUG-2003 (revised)
DT 06-APR-1998 (first entry)
XX
XX
DE Rat type G protein-coupled receptor ligand fragment 5.

KW G protein-coupled receptor; ligand binding; pharmaceutical; modulator;
KW pituitary; central nervous system; pancreas; propylactic;
KW therapeutic agent.

XX		
XX		
FN	W09724436-A2.	
XX		
PD	10-JUL-1997.	
XX		
PF	26-DEC-1996;	96WO-JP003821.
XX		
PR	28-DEC-1995;	95JP-00343371.
PR	15-MAR-1996;	96JP-00059419.
PR	12-AUG-1996;	96JP-00211805.
PR	18-SEP-1996;	96JP-00246573.
XX		
PA	(TAKE) TAKEDA CHEM IND LTD.	

FA	(PAGE) LARSENDA CHERM AND LTD.
XX	
PI	Hinuma S, Habata Y, Kawamata Y, Hosoya M, Fujii R, Fukusumi S;
PI	Kitada C;
XX	
XX	WPI; 1997-363672/33.
DR	N-PSDB; AAV02425.
DR	
XX	
XX	Ligand peptide for G protein-coupled receptor - acts by modulating
PT	function in the central nervous system, pancreas and pituitary gland.
PT	
XX	
XX	Claim 2; Page 180; 258pp; English.
FS	

This sequence represents a peptide fragment from a novel rat type I ligand polypeptide corresponding to amino acid residues 33 to 53 of the sequence represented in AAW1381 and is used in an assay to monitor ligand binding to the G protein-coupled receptor protein. Pharmaceutical compositions containing this ligand may be used as a pituitary function modulator, a central nervous system modulator or a pancreatic function modulator. This ligand could have specific applications as a prophylactic or therapeutic agent for dementia, depression, hyperkinetic syndrome, disturbance of

CC consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone
 CC secretory disease, hyper- and polyphagia, hyperlipidaemia, hyperprolactinaemia, diabetes,
 CC hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,
 CC amyotrophic lateral sclerosis, acute myocardial infarction, infertility,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis and/or oligogalactia. Assays can also be developed to screen
 CC compounds which are capable of altering the binding activity of the
 CC ligand affecting activation of the G protein-coupled receptor protein.
 CC (Updated on 27-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 21 AA;

Query Match 100.0%; Score 79; DB 2; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.5e-06; Mismatches 0; Indels 0; Gaps 0;
 Matches 13; Conservative 0;

QY 1 TPDINPAWYTGKG 13
 |||||
 DB 1 TPDINPAWYTGKG 13

RESULT 11
 AAB10359
 ID AAB10359 standard; peptide; 21 AA.

XX
 AC AAB10359;

DT 24-NOV-2000 (first entry)

DE Rat oxytocin secretion promoting peptide SEQ ID NO: 22.

KW Rat; oxytocin secretion promoter; G protein-coupled receptor protein;
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;
 KW veterinary medicine; milk production.

OS Rattus sp.

PN WO200038704-A1.

PD 06-JUL-2000.

PF 22-DEC-1999; 99NO-JP007199.

PR 25-DEC-1998; 98JP-00369585.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Matsumoto H, Kitada C, Hinuma S;

XX WPI; 2000-452298/39.

XX Physiologically-active polypeptide recognized as ligand by G protein-
 PT coupled receptor protein, for promoting secretion of oxytocin, as drugs
 PT for diseases relating to oxytocin secretion and in veterinary medicine.

PS Disclosure; Page 58; 72pp; Japanese.

XX This invention describes a novel oxytocin secretion-regulating agent
 CC which contains a ligand peptide or its salt for the G protein-coupled
 CC receptor protein. It is useful in the form of drugs for ameliorating,
 CC preventing and treating diseases relating to oxytocin secretion e.g. weak
 CC pains and atonic bleeding, before and after expulsion of placenta,
 CC uterine recovery failure, caesarean section, stoppage of artificial
 CC fertilization or galactostasis and is also applicable in veterinary
 CC medicine for promoting milk production in cow, goat and pig. This
 CC sequence represents a rat peptide which acts as an oxytocin secretion
 CC promoter

XX Sequence 21 AA;

Query Match 100.0%; Score 79; DB 3; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.5e-06; Mismatches 0; Indels 0; Gaps 0;
 Matches 13; Conservative 0;

QY 1 TPDINPAWYTGKG 13
 |||||
 DB 1 TPDINPAWYTGKG 13

RESULT 12

AAG62528

ID AAG62528 standard; peptide; 21 AA.

XX
 AC AAG62528;

DT 24-AUG-2001 (first entry)

DE Rat CRH releasing protein related peptide SEQ ID NO: 22.

KW Rat; corticotrophin releasing hormone; CRH; G protein receptor ligand;
 KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;
 KW Addison's disease; adrenal gland hyperfunction; obesity.

OS Rattus sp.

PN WO200135984-A1.

PD 25-MAY-2001.

PF 17-NOV-2000; 2000WO-JP008119.

PR 18-NOV-1999; 99JP-00327900.

PR 26-SEP-2000; 2000JP-00297073.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Kitada C, Matsumoto H, Hinuma S;

XX WPI; 2001-355552/37.

XX Use of G protein receptor ligand or peptide for controlling corticotropin
 PT releasing hormone secretion.

PS Disclosure; Page 70; 90pp; Japanese.

XX The present sequence describes a method of controlling the secretion of
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein
 CC receptor ligand. This can be used to control the secretion of CRH and is
 CC useful as an analgesic or for treating, preventing or ameliorating
 CC diseases associated with CRH secretion such as hyperaldosteronism,
 CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair
 CC loss, and hypotension), adrenal gland hyperfunction and obesity. The
 CC present sequence is a peptide used in the exemplification of the
 CC invention

XX Sequence 21 AA;

Query Match 100.0%; Score 79; DB 4; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.5e-06; Mismatches 0; Indels 0; Gaps 0;
 Matches 13; Conservative 0;

QY 1 TPDINPAWYTGKG 13
 |||||
 DB 1 TPDINPAWYTGKG 13

RESULT 13

ABU60841

ID ABU60841 standard; peptide; 21 AA.

XX
 AC ABU60841;

XX

DT	06-MAY-2003	(first entry)	
XX	Peptide production by gene recombination associated peptide #25.		
XX	Peptide production; low-molecular peptide; KiSS-1; GPR8 ligand;		
KW	gene recombination.		
XX	Rattus sp.		
OS			
XX	WO200292829-A1.		
PN			
XX	21-NOV-2002.		
XX			
PF	16-MAY-2002; 2002WO-JP004735.		
XX			
PR	17-MAY-2001; 2001JP-00147341.		
XX	(TAKE) TAKEDA CHEM IND LTD.		
PA			
XX	Nishimura O, Suenaga M, Ito T, Kitada C;		
PI			
XX	WPI; 2003-129302/12.		
DR			
XX	Process for producing peptides e.g. KiSS-1 peptide and GPR8 ligand for		
PT	subsequent applications by gene recombination technique through tandem		
PT	repeats to provide precursor protein with specific cleavage sites.		
XX			
PS	Disclosure; Page 66; 87pp; Japanese.		
XX			
CC	The invention describes a method of producing a peptide comprising the		
CC	excision of the N and C-terminals of a target peptide with enzymes or		
CC	chemically through the attached cleavage sites repeated by ligation in a		
CC	precursor protein. The method is for producing (low-molecular) peptides		
CC	e.g. KiSS-1 peptide and GPR8 ligand for subsequent applications by the		
CC	gene recombination technique through tandem repeats to provide a		
CC	precursor protein with specific cleavage sites. With this method, peptide		
CC	production can be carried out easily to provide large quantities of the		
CC	required peptides. This is the amino acid sequence of a peptide		
CC	associated with the peptide production method of the invention		
XX			
SQ	Sequence 21 AA;		
	Query Match	100.0%; Score 79; DB 6; Length 21;	
	Best Local Similarity	100.0%; Pred. No. 4.5e-06;	
	Matches	13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	TPDINPAWYTGKG 13	
Db	1	TPDINPAWYTGKG 13	
RESULT 14			
AAW31389			
ID	AAW31389	standard; peptide; 22 AA.	
XX			
AC	AAW31389;		
XX			
DT	27-AUG-2003	(revised)	
DT	06-APR-1998	(first entry)	
XX			
DE	Rat type G	protein-coupled receptor ligand fragment 6.	
XX			
KW	G	protein-coupled receptor; ligand binding; pharmaceutical; modulator;	
KW	pituitary; central nervous system; pancreas; prophylactic;		
XX	therapeutic agent.		
XX			
OS	Rattus sp.		
XX			
FN	WO9724436-A2.		
XX			
PD	10-JUL-1997.		
XX			
PF	26-DEC-1996;	96WO-JP003821.	

```

PR 25-DEC-1998; 98JP-00369585.
XX (TAKE ) TAKEDA CHEM IND LTD.
XX
XX PI Matsumoto H, Kitada C, Hinuma S;
XX DR WPI; 2000-452298/39.
XX
XX PT Physiologically-active polypeptide recognized as ligand by G protein-
XX PT coupled receptor protein, for promoting secretion of oxytocin, as drugs
XX PT for diseases relating to oxytocin secretion and in veterinary medicine.
XX PS Disclosure; Page 59; 72pp; Japanese.
XX
XX CC This invention describes a novel oxytocin secretion-regulating agent
XX CC which contains a ligand peptide or its salt for the G protein-coupled
XX CC receptor protein. It is useful in the form of drugs for ameliorating,
XX CC preventing and treating diseases relating to oxytocin secretion e.g. weak
XX CC pains and atonic bleeding, before and after expulsion of placenta,
XX CC uterine recovery failure, caesarean section, stoppage of artificial
XX CC fertilization or galactostasis and is also applicable in veterinary
XX CC medicine for promoting milk production in cow, goat and pig. This
XX CC sequence represents a rat peptide which acts as an oxytocin secretion
XX CC promoter
XX SQ Sequence 22 AA;

Query Match 100.0%; Score 79; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 4.7e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGK 13
   |||||
Db 1 TPDINPAWYTGK 13

Search completed: August 12, 2004, 14:43:55
Job time : 27.7122 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:49:10 ; Search time 22.9767 Seconds
(without alignments)
177.617 Million cell updates/sec

Title: US-09-700-643a-3_COPY_12_24

Perfect score: 79

Sequence: 1 TPDINPAWYTGK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1292805 segs, 313927144 residues

Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	20	9	US-09-932-161-17
2	79	100.0	20	13	US-10-044-592-6
3	79	100.0	20	14	US-10-096-777-17
4	79	100.0	31	9	US-09-932-161-14
5	79	100.0	31	13	US-10-044-592-4
6	79	100.0	31	13	US-10-044-592-5
7	79	100.0	31	14	US-10-096-777-14
8	79	100.0	70	13	US-10-044-592-90
9	79	100.0	82	13	US-10-044-592-1
10	79	100.0	86	13	US-10-044-592-96
11	79	100.0	91	13	US-10-044-592-94
12	79	100.0	19	13	US-10-044-592-27
13	74	93.7	20	9	US-09-932-161-16
14	74	93.7	20	13	US-10-044-592-42
15	74	93.7	20	14	US-10-096-777-16

Sequence 43, Appl
Sequence 44, Appl
Sequence 78, Appl
Sequence 26, Appl
Sequence 13, Appl
Sequence 39, Appl
Sequence 13, Appl
Sequence 40, Appl
Sequence 41, Appl
Sequence 28, Appl
Sequence 38, Appl
Sequence 82, Appl
Sequence 84, Appl
Sequence 86, Appl
Sequence 88, Appl
Sequence 18, Appl
Sequence 18, Appl
Sequence 15, Appl
Sequence 55, Appl
Sequence 92, Appl
Sequence 8, Appl
Sequence 268092,
Sequence 197, App
Sequence 6327, Ap
Sequence 535, App
Sequence 535, App
Sequence 574, App
Sequence 1690, Ap
Sequence 74, Appl
Sequence 356, App

ALIGNMENTS

RESULT 1

US-09-932-161-17
; Sequence 17, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; FILE REFERENCE P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Rattus
US-09-932-161-17

Query Match 100.0%; Score 79; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPDINPAWYTGK 13

Db 1 TPDINPAWYTGK 13

RESULT 2

US-10-044-592-6
; Sequence 6, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji

```

; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 6
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(20)
; OTHER INFORMATION: antigen
;
US-10-044-592-6

```

```

Query Match      100.0%; Score 79; DB 13; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 TPDINPAWYTGK 13
Db      1 TPDINPAWYTGK 13

```

```

RESULT 3
US-10-096-777-17
; Sequence 17, Application US/10096777
; Publication No. US20030171270A1
; GENERAL INFORMATION:
; APPLICANT: Civealli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; TITLE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)
; FILE REFERENCE: PUC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Rattus
;
US-10-096-777-17

```

```

Query Match      100.0%; Score 79; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 TPDINPAWYTGK 13
Db      1 TPDINPAWYTGK 13

```

```

RESULT 4
US-09-932-161-14
; Sequence 14, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Civealli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; TITLE OF INVENTION: Promoting Wakefulness and Sleep
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161

```

```

; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Rattus
;
US-09-932-161-14

```

```

Query Match      100.0%; Score 79; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 TPDINPAWYTGK 13
Db      12 TPDINPAWYTGK 24

```

```

RESULT 5
US-10-044-592-4
; Sequence 4, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 4
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Murine
;
US-10-044-592-4

```

```

Query Match      100.0%; Score 79; DB 13; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 TPDINPAWYTGK 13
Db      12 TPDINPAWYTGK 24

```

```

RESULT 6
US-10-044-592-5
; Sequence 5, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Fukusumi, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:

```

```
; SEQ ID NO 5
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(31)
; OTHER INFORMATION: antigen
US-10-044-592-5

Query Match      100.0%; Score 79; DB 13; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGK 13
Db 12 TPDINPAWYTGK 24

RESULT 7
US-10-096-777-14
; Sequence 14, Application US/10096777
; Publication No. US20030171270A1
; GENERAL INFORMATION:
; APPLICANT: Civelii, Olivier
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; TITLE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Rattus
US-10-096-777-14

Query Match      100.0%; Score 79; DB 14; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGK 13
Db 12 TPDINPAWYTGK 24

RESULT 8
US-10-044-592-90
; Sequence 90, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 90
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Rat

Query Match      100.0%; Score 79; DB 13; Length 70;
Best Local Similarity 100.0%; Pred. No. 5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGK 13
Db 33 TPDINPAWYTGK 45

US-10-044-592-90

Query Match      100.0%; Score 79; DB 13; Length 70;
Best Local Similarity 100.0%; Pred. No. 5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGK 13
Db 33 TPDINPAWYTGK 45

RESULT 9
US-10-044-592-1
; Sequence 1, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 1
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Murine
US-10-044-592-1

Query Match      100.0%; Score 79; DB 13; Length 82;
Best Local Similarity 100.0%; Pred. No. 5.8e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGK 13
Db 32 TPDINPAWYTGK 44

RESULT 10
US-10-044-592-96
; Sequence 96, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 96
; LENGTH: 86
; TYPE: PRT
; ORGANISM: mammalian
; NAME/KEY: misc feature
; LOCATION: (1224)..(1243)
; OTHER INFORMATION: Bracket region depicted in FIG 39.
US-10-044-592-96
```

Query Match 100.0%; Score 79; DB 13; Length 86;
Best Local Similarity 100.0%; Pred. No. 6.1e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGK 13
Db 33 TPDINPAWYTGK 45

RESULT 11

US-10-044-592-94
; Sequence 94, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 94
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Mammalian
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(31)
; OTHER INFORMATION: primer
; NAME/KEY: misc feature
; LOCATION: (925)-(955)
; OTHER INFORMATION: primer
; NAME/KEY: misc feature
; LOCATION: (1)-(955)
; OTHER INFORMATION: insert fragment of pmGB3 as depicted in FIG 36 and 37.
US-10-044-592-94

Query Match 100.0%; Score 79; DB 13; Length 91;
Best Local Similarity 100.0%; Pred. No. 6.4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGK 13
Db 33 TPDINPAWYTGK 45

RESULT 12

US-10-044-592-27
; Sequence 27, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:

; SEQ ID NO 27
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-27

Query Match 93.7%; Score 74; DB 13; Length 19;
Best Local Similarity 92.3%; Pred. No. 8.8e-05;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGK 13
Db 1 TPDINPAWYAGRG 13

RESULT 13

US-09-932-161-16
; Sequence 16, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Civelis, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; TITLE OF INVENTION: Promoting Wakefulness and Sleep
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-932-161-16

Query Match 93.7%; Score 74; DB 9; Length 20;
Best Local Similarity 92.3%; Pred. No. 9.2e-05;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGK 13
Db 1 TPDINPAWYAGRG 13

RESULT 14

US-10-044-592-42
; Sequence 42, Application US/10044592
; Publication No. US20020143152A1
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shoji
; TITLE OF INVENTION: Polypeptides, their Production and Use
; FILE REFERENCE: 2463US2P
; CURRENT APPLICATION NUMBER: US/10/044,592
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/403639
; PRIOR FILING DATE: 1999-25-10
; PRIOR APPLICATION NUMBER: PCT/JP98/01923
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: JP 9-109974
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE:
; SEQ ID NO 42
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Bovine
US-10-044-592-42

Query Match 93.7%; Score 74; DB 13; Length 20;
Best Local Similarity 92.3%; Pred. No. 9.2e-05;

Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TPDINPAWYTGK 13
| | | | | | | | | |
Db 1 TPDINPAWYAGK 13

RESULT 15
US-10-096-777-16
; Sequence 16, Application US/10096777
; Publication No. US20030171270A1
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; TITLE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Bos taurus
US-10-096-777-16

Query Match 93.7%; Score 74; DB 14; Length 20;
Best Local Similarity 92.3%; Pred. NO. 9.2e-05;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TPDINPAWYTGK 13
| | | | | | | | | |
Db 1 TPDINPAWYAGK 13

Search completed: August 12, 2004, 15:22:51
Job time : 23.9767 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:45:00 ; Search time 11.7151 Seconds
(without alignments)
130.133 Million cell updates/sec

Title: US-09-700-643A-3_COPY_12_24
Perfect score: 79
Sequence: 1 TPDINPAWYTGKG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 650653 seqs, 11721403 residues

Total number of hits satisfying chosen parameters: 650653

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New.*
1: /cgn2_6/prodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/prodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/prodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/prodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/prodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/prodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/prodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES										
Result No.	Score	Query Match %	Length	DB ID	Description					
1	79	100.0	20	6	US-10-477-712B-35	Sequence 35, Appl				
2	79	100.0	20	6	US-10-477-712B-84	Sequence 84, Appl				
3	79	100.0	21	6	US-10-477-712B-36	Sequence 36, Appl				
4	79	100.0	21	6	US-10-477-712B-76	Sequence 76, Appl				
5	79	100.0	22	6	US-10-477-712B-37	Sequence 37, Appl				
6	79	100.0	22	6	US-10-477-712B-80	Sequence 80, Appl				
7	79	100.0	31	6	US-10-477-712B-21	Sequence 21, Appl				
8	79	100.0	31	6	US-10-477-712B-32	Sequence 32, Appl				
9	79	100.0	32	6	US-10-477-712B-33	Sequence 33, Appl				
10	79	100.0	33	6	US-10-477-712B-34	Sequence 34, Appl				
11	75	94.9	21	6	US-10-669-228B-15	Sequence 15, Appl				
12	74	93.7	19	6	US-10-477-712B-25	Sequence 25, Appl				
13	74	93.7	20	6	US-10-477-712B-29	Sequence 29, Appl				
14	74	93.7	20	6	US-10-477-712B-82	Sequence 82, Appl				
15	74	93.7	21	6	US-10-477-712B-23	Sequence 23, Appl				
16	74	93.7	21	6	US-10-477-712B-30	Sequence 30, Appl				
17	74	93.7	22	6	US-10-477-712B-31	Sequence 31, Appl				
18	74	93.7	22	6	US-10-477-712B-78	Sequence 78, Appl				
19	74	93.7	22	6	US-10-477-712B-24	Sequence 24, Appl				
20	74	93.7	31	6	US-10-477-712B-20	Sequence 20, Appl				
21	74	93.7	31	6	US-10-477-712B-26	Sequence 26, Appl				
22	74	93.7	32	6	US-10-477-712B-27	Sequence 27, Appl				
23	74	93.7	32	6	US-10-477-712B-28	Sequence 28, Appl				
24	73	92.4	20	6	US-10-477-712B-85	Sequence 85, Appl				
25	73	92.4	21	6	US-10-477-712B-77	Sequence 77, Appl				
26	71	89.9	22	6	US-10-477-712B-81	Sequence 81, Appl				

Sequence 41, Appl
Sequence 83, Appl
Sequence 42, Appl
Sequence 75, Appl
Sequence 43, Appl
Sequence 79, Appl
Sequence 22, Appl
Sequence 38, Appl
Sequence 39, Appl
Sequence 40, Appl
Sequence 18159, A
Sequence 143, App
Sequence 143, App
Sequence 28900, A
Sequence 340709,
Sequence 3867, Ap
Sequence 3649, Ap
Sequence 3866, Ap

ALIGNMENTS

RESULT 1
US-10-477-712B-35
; Sequence 35, Application US/10477712B
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: A Method for Producing A Peptide
; FILE REFERENCE: OKA-0213
; CURRENT APPLICATION NUMBER: US/10477,712B
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: JP 2001-147341
; PRIOR FILING DATE: 2001-05-17
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 35
; LENGTH: 20
; TYPE: PPT
; ORGANISM: Rat
US-10-477-712B-35

Query Match 100.0%; Score 79; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.5e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGKG 13
Db 1 TPDINPAWYTGKG 13

RESULT 2
US-10-477-712B-84
; Sequence 84, Application US/10477712B
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: A Method for Producing A Peptide
; FILE REFERENCE: OKA-0213
; CURRENT APPLICATION NUMBER: US/10477,712B
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: JP 2001-147341
; PRIOR FILING DATE: 2001-05-17
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 84
; LENGTH: 20
; TYPE: PPT
; ORGANISM: Bovine
US-10-477-712B-84

Query Match 100.0%; Score 79; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.5e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYTGK 13
Db 1 TPDINPAWYTGK 13

RESULT 3
US-10-477-712B-36
; Sequence 36, Application US/10477712B
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: A Method for Producing A Peptide
; FILE REFERENCE: OKA-0213
; CURRENT APPLICATION NUMBER: US/10/477,712B
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: JP 2001-147341
; PRIOR FILING DATE: 2001-05-17
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 36
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Rat
US-10-477-712B-36

Query Match 100.0%; Score 79; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYTGK 13
Db 1 TPDINPAWYTGK 13

RESULT 4
US-10-477-712B-76
; Sequence 76, Application US/10477712B
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: A Method for Producing A Peptide
; FILE REFERENCE: OKA-0213
; CURRENT APPLICATION NUMBER: US/10/477,712B
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: JP 2001-147341
; PRIOR FILING DATE: 2001-05-17
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 76
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Bovine
US-10-477-712B-76

Query Match 100.0%; Score 79; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYTGK 13
Db 1 TPDINPAWYTGK 13

RESULT 5
US-10-477-712B-37
; Sequence 37, Application US/10477712B
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: A Method for Producing A Peptide
; FILE REFERENCE: OKA-0213
; CURRENT APPLICATION NUMBER: US/10/477,712B
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: JP 2001-147341
; PRIOR FILING DATE: 2001-05-17
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 37
; LENGTH: 22

; TYPE: PRT
; ORGANISM: Human
US-10-477-712B-37

Query Match 100.0%; Score 79; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYTGK 13
Db 1 TPDINPAWYTGK 13

RESULT 6
US-10-477-712B-80
; Sequence 80, Application US/10477712B
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: A Method for Producing A Peptide
; FILE REFERENCE: OKA-0213
; CURRENT APPLICATION NUMBER: US/10/477,712B
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: JP 2001-147341
; PRIOR FILING DATE: 2001-05-17
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 80
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Bovine
US-10-477-712B-80

Query Match 100.0%; Score 79; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYTGK 13
Db 1 TPDINPAWYTGK 13

RESULT 7
US-10-477-712B-21
; Sequence 21, Application US/10477712B
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: A Method for Producing A Peptide
; FILE REFERENCE: OKA-0213
; CURRENT APPLICATION NUMBER: US/10/477,712B
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: JP 2001-147341
; PRIOR FILING DATE: 2001-05-17
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 21
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Rat
US-10-477-712B-21

Query Match 100.0%; Score 79; DB 6; Length 31;
Best Local Similarity 100.0%; Pred. No. 3.9e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYTGK 13
Db 12 TPDINPAWYTGK 24

RESULT 8
US-10-477-712B-32
; Sequence 32, Application US/10477712B
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: A Method for Producing A Peptide

```
; FILE REFERENCE: OKA-0213
; CURRENT APPLICATION NUMBER: US/10/477,712B
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: JP 2001-147341
; PRIOR FILING DATE: 2001-05-17
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 32
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Rat
US-10-477-712B-32

Query Match      100.0%; Score 79; DB 6; Length 31;
Best Local Similarity 100.0%; Pred. No. 3.9e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGKG 13
Db 12 TPDINPAWYTGKG 24

RESULT 9
US-10-477-712B-33
; Sequence 33, Application US/10477712B
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: A Method for Producing A Peptide
; FILE REFERENCE: OKA-0213
; CURRENT APPLICATION NUMBER: US/10/477,712B
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: JP 2001-147341
; PRIOR FILING DATE: 2001-05-17
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 33
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Rat
US-10-477-712B-33

Query Match      100.0%; Score 79; DB 6; Length 32;
Best Local Similarity 100.0%; Pred. No. 4e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGKG 13
Db 12 TPDINPAWYTGKG 24

RESULT 10
US-10-477-712B-34
; Sequence 34, Application US/10477712B
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: A Method for Producing A Peptide
; FILE REFERENCE: OKA-0213
; CURRENT APPLICATION NUMBER: US/10/477,712B
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: JP 2001-147341
; PRIOR FILING DATE: 2001-05-17
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 34
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Rat
US-10-477-712B-34

Query Match      100.0%; Score 79; DB 6; Length 33;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGKG 13
Db 12 TPDINPAWYTGKG 24

; FILE REFERENCE: OKA-0213
; CURRENT APPLICATION NUMBER: US/10/477,712B
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: JP 2001-147341
; PRIOR FILING DATE: 2001-05-17
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 15
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide
; OTHER INFORMATION: the C-terminus of the polypeptide is amide (-CONH2) form
US-10-069-228B-15

Query Match      94.9%; Score 75; DB 6; Length 31;
Best Local Similarity 92.3%; Pred. No. 1.8e-05;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGKG 13
Db 12 TPDINPAWYTGKG 24

RESULT 12
US-10-477-712B-25
; Sequence 25, Application US/10477712B
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: A Method for Producing A Peptide
; FILE REFERENCE: OKA-0213
; CURRENT APPLICATION NUMBER: US/10/477,712B
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: JP 2001-147341
; PRIOR FILING DATE: 2001-05-17
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 25
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Bovine
US-10-477-712B-25

Query Match      93.7%; Score 74; DB 6; Length 19;
Best Local Similarity 92.3%; Pred. No. 1.6e-05;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGKG 13
Db 1 TPDINPAWYAGRG 13

RESULT 13
US-10-477-712B-29
; Sequence 29, Application US/10477712B
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: A Method for Producing A Peptide
; FILE REFERENCE: OKA-0213
; CURRENT APPLICATION NUMBER: US/10/477,712B
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: JP 2001-147341
```

; PRIOR FILING DATE: 2001-05-17
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 29
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Bovine
US-10-477-712B-29

Query Match 93.7%; Score 74; DB 6; Length 20;
Best Local Similarity 92.3%; Pred. No. 1.7e-05;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPDINPAWYTG RG 13
| | | | | | | | | |
Db 1 TPDINPAWYAG RG 13

RESULT 14
US-10-477-712B-82
; Sequence 82, Application US/10477712B
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: A Method for Producing A Peptide
; FILE REFERENCE: OKA-0213
; CURRENT APPLICATION NUMBER: US/10/477,712B
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: JP 2001-147341
; PRIOR FILING DATE: 2001-05-17
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 82
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Bovine
US-10-477-712B-82

Query Match 93.7%; Score 74; DB 6; Length 20;
Best Local Similarity 92.3%; Pred. No. 1.7e-05;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPDINPAWYTG RG 13
| | | | | | | | | |
Db 1 TPDINPAWYAG RG 13

RESULT 15
US-10-477-712B-23
; Sequence 23, Application US/10477712B
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: A Method for Producing A Peptide
; FILE REFERENCE: OKA-0213
; CURRENT APPLICATION NUMBER: US/10/477,712B
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: JP 2001-147341
; PRIOR FILING DATE: 2001-05-17
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 23
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Bovine
US-10-477-712B-23

Query Match 93.7%; Score 74; DB 6; Length 21;
Best Local Similarity 92.3%; Pred. No. 1.8e-05;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPDINPAWYTG RG 13
| | | | | | | | | |
Db 1 TPDINPAWYAG RG 13

Search completed: August 12, 2004, 15:17:35
Job time : 11.7151 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 14:37:36 ; Search time 7.40698 Seconds
(without alignments)
90.609 Million cell updates/sec

Title: US-09-700-643A-3_COPY_12_24

Perfect score: 79

Sequence: 1 TPDINPAWYTGSG 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A COMB.pdp.*
- 2: /cgn2_6/ptodata/2/iaa/5B COMB.pdp.*
- 3: /cgn2_6/ptodata/2/iaa/5A COMB.pdp.*
- 4: /cgn2_6/ptodata/2/iaa/5B COMB.pdp.*
- 5: /cgn2_6/ptodata/2/iaa/5A COMB.pdp.*
- 6: /cgn2_6/ptodata/2/iaa/5B COMB.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	20	3	US-09-105-678A-40
2	79	100.0	20	3	US-08-776-971-50
3	79	100.0	20	3	US-09-421-208-40
4	79	100.0	20	4	US-09-560-915-17
5	79	100.0	21	3	US-09-105-678A-41
6	79	100.0	21	3	US-08-776-971-51
7	79	100.0	21	3	US-09-421-208-41
8	79	100.0	22	3	US-09-105-678A-42
9	79	100.0	22	3	US-08-776-971-52
10	79	100.0	22	3	US-09-421-208-42
11	79	100.0	31	3	US-09-105-678A-8
12	79	100.0	31	3	US-09-105-678A-37
13	79	100.0	31	3	US-09-172-353-4
14	79	100.0	31	3	US-08-776-971-47
15	79	100.0	31	3	US-09-421-208-8
16	79	100.0	31	3	US-09-421-208-37
17	79	100.0	31	4	US-09-560-915-14
18	79	100.0	31	4	US-09-799-955-4
19	79	100.0	32	3	US-09-105-678A-38
20	79	100.0	32	3	US-08-776-971-48
21	79	100.0	32	3	US-09-421-208-38
22	79	100.0	33	3	US-09-105-678A-39
23	79	100.0	33	3	US-08-776-971-49
24	79	100.0	33	3	US-09-421-208-39
25	79	100.0	83	3	US-08-776-971-45
26	79	100.0	83	3	US-08-776-971-124
27	79	100.0	83	3	US-08-776-971-137

28	74	93.7	19	3	US-09-105-678A-30	Sequence 30, Appl
29	74	93.7	19	3	US-08-776-971-4	Sequence 4, Appl
30	74	93.7	19	3	US-09-421-208-30	Sequence 30, Appl
31	74	93.7	20	3	US-09-105-678A-34	Sequence 34, Appl
32	74	93.7	20	3	US-08-776-971-8	Sequence 8, Appl
33	74	93.7	20	3	US-08-776-971-98	Sequence 98, Appl
34	74	93.7	20	3	US-09-421-208-34	Sequence 34, Appl
35	74	93.7	20	4	US-09-560-915-16	Sequence 16, Appl
36	74	93.7	21	3	US-09-105-678A-35	Sequence 35, Appl
37	74	93.7	21	3	US-08-776-971-9	Sequence 9, Appl
38	74	93.7	21	3	US-09-421-208-35	Sequence 35, Appl
39	74	93.7	22	3	US-09-105-678A-36	Sequence 36, Appl
40	74	93.7	22	3	US-08-776-971-10	Sequence 10, Appl
41	74	93.7	22	3	US-09-421-208-36	Sequence 36, Appl
42	74	93.7	25	3	US-08-776-971-111	Sequence 111, Appl
43	74	93.7	29	3	US-09-105-678A-29	Sequence 29, Appl
44	74	93.7	29	3	US-08-776-971-3	Sequence 3, Appl
45	74	93.7	29	3	US-09-421-208-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1
US-09-105-678A-40
; Sequence 40, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-105-678A-40

Query Match 100.0%; Score 79; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TPDINPAWYTGSG 13

Db 1 TPDINPAWYTG RG 13
|||||
RESULT 2
US-08-776-971-50
; Sequence 50, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; Habata, Yugo
; Kawamata, Yuji
; Hosoya, Masaki
; Fujii, Ryo
; Fukusumi, Shoji
; Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 50:
US-08-776-971-50
Query Match 100.0%; Score 79; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 TPDINPAWYTG RG 13
Db 1 TPDINPAWYTG RG 13
|||||
RESULT 3
US-08-776-971-50
Query Match 100.0%; Score 79; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 TPDINPAWYTG RG 13
Db 1 TPDINPAWYTG RG 13
|||||
US-09-421-208-40
; Sequence 40, Application US/09421208
; Patent No. 6259561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-421-208-40
Query Match 100.0%; Score 79; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 TPDINPAWYTG RG 13
Db 1 TPDINPAWYTG RG 13
|||||
RESULT 4
US-09-560-915-17
; Sequence 17, Application US/09560915
; Patent No. 6383764
; GENERAL INFORMATION:
; APPLICANT: Civeilli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; TITLE OF INVENTION: Relating To Prolactin Releasing Peptide (PzRP)
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/09/560,915
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Rattus

US-09-560-915-17

Query Match 100.0%; Score 79; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGKG 13

Db 1 TPDINPAWYTGKG 13

RESULT 5

US-09-105-678A-41
; Sequence 41, Application US/09105678A

; Patent No. 6103882

; GENERAL INFORMATION:

; APPLICANT: Suenaga, Masato

; APPLICANT: Moriya, Takeo

; APPLICANT: Tanaka, Yoko

; APPLICANT: Nishimura, Osamu

; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND

; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

; STREET: 130 Water Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/105,678A

; FILING DATE: 28-JUN-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 172118/1997

; FILING DATE: 27-JUN-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Conlin, David G.

; REGISTRATION NUMBER: 27,026

; REFERENCE/DOCKET NUMBER: 48466-342

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-523-3400

; TELEFAX: 617-523-6440

; INFORMATION FOR SEQ ID NO: 41:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 21 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-09-105-678A-41

Query Match 100.0%; Score 79; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGKG 13

Db 1 TPDINPAWYTGKG 13

RESULT 6

US-08-776-971-51

; Sequence 51, Application US/08776971B

; Patent No. 6228984

; GENERAL INFORMATION:

; APPLICANT: Hinuma, Shuji

; APPLICANT: Habata, Yugo

; APPLICANT: Kawamata, Yuji

; APPLICANT: KAWAMATA, YUJI

; APPLICANT: KAWAMATA, YUJI

; APPLICANT: KAWAMATA, YUJI

; APPLICANT: KAWAMATA, YUJI

; APPLICANT: KAWAMATA, YUJI

; APPLICANT: KAWAMATA, YUJI

; APPLICANT: KAWAMATA, YUJI

; APPLICANT: KAWAMATA, YUJI

Hosoya, Masaki

Fujii, Ryo

Fukusumi, Shoji

Kitada, Chieko

TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE

NUMBER OF SEQUENCES: 140

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

MEDIUM TYPE: Diskette

COMPUTER: IBM compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/776,971B

FILING DATE: 06-Feb-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP96/03821

FILING DATE: 28-DEC-1996

APPLICATION NUMBER: JP 7/343371

FILING DATE: 28-DEC-1995

APPLICATION NUMBER: JP 8/59419

FILING DATE: 15-MAR-1996

APPLICATION NUMBER: JP 8/211805

FILING DATE: 12-AUG-1996

APPLICATION NUMBER: JP 8/246573

FILING DATE: 18-SEP-1996

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 47176

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 51:

SEQUENCE CHARACTERISTICS:

LENGTH: 21 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: internal

SEQUENCE DESCRIPTION: SEQ ID NO: 51:

US-08-776-971-51

Query Match 100.0%; Score 79; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGKG 13

Db 1 TPDINPAWYTGKG 13

RESULT 7

US-09-421-208-41

; Sequence 41, Application US/09421208

; Patent No. 6258561

; GENERAL INFORMATION:

; APPLICANT: Suenaga, Masato

; APPLICANT: Moriya, Takeo

; APPLICANT: Tanaka, Yoko

; APPLICANT: Nishimura, Osamu

; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND

; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/421,208
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/105,678
FILING DATE: 26-JUN-1998
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-421-208-41

Query Match 100.0%; Score 79; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0;

Qy 1 TPDINPAWYTGK 13
Db 1 TPDINPAWYTGK 13

RESULT 8
US-09-105-678A-42
Sequence 42, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-42
Query Match 100.0%; Score 79; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TPDINPAWYTGK 13
Db 1 TPDINPAWYTGK 13
RESULT 9
US-08-776-971-52
Sequence 52, Application US/08776971B
Patent No. 6228984
GENERAL INFORMATION:
APPLICANT: Hinuma, Shuji
APPLICANT: Habata, Yugo
APPLICANT: Kawamata, Yuji
APPLICANT: Hosoya, Masaki
APPLICANT: Fujii, Ryo
APPLICANT: Fukusumi, Shoji
APPLICANT: Kitada, Chieko
TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,971B
FILING DATE: 06-Feb-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03821
FILING DATE: 28-DEC-1996
APPLICATION NUMBER: JP 7/343371
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: JP 8/59419
FILING DATE: 15-MAR-1996
APPLICATION NUMBER: JP 8/211805
FILING DATE: 12-AUG-1996
APPLICATION NUMBER: JP 8/246573
FILING DATE: 18-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 47176
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 52:

SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-08-776-971-52

Query Match 100.0%; Score 79; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGGRG 13
Db 1 TPDINPAWYTGGRG 13

RESULT 10
US-09-421-208-42
; Sequence 42, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE: 27-JUN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-421-208-42

Query Match 100.0%; Score 79; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGGRG 13
Db 1 TPDINPAWYTGGRG 13

SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-08-776-971-52

QY 1 TPDINPAWYTGGRG 13
Db 1 TPDINPAWYTGGRG 24

RESULT 11
US-09-105-678A-8
; Sequence 8, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-105-678A-8

Query Match 100.0%; Score 79; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGGRG 13
Db 12 TPDINPAWYTGGRG 24

RESULT 12
US-09-105-678A-37
; Sequence 37, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA

Query Match 100.0%; Score 79; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGGRG 13
Db 1 TPDINPAWYTGGRG 13

```

; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-105-678A-37

Query Match 100.0%; Score 79; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGK 13
Db 12 TPDINPAWYTGK 24

RESULT 13
US-09-172-353-4
; Sequence 4, Application US/09172353
; Patent No. 6197530
; GENERAL INFORMATION:
; APPLICANT: Stricker-Kongra, Alain
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: GPR10 AS A TARGET FOR IDENTIFYING WEIGHT MODULATING COMPOUNDS
; FILE REFERENCE: 07334/102001
; CURRENT APPLICATION NUMBER: US/09/172,353
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 31
; TYPE: PPT
; ORGANISM: Mus musculus
; US-09-172-353-4

Query Match 100.0%; Score 79; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGK 13
Db 12 TPDINPAWYTGK 24

RESULT 14
US-08-776-971-47
; Sequence 47, Application US/0876971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Habata, Yugo
; APPLICANT: Kawamata, Yuji
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

; Hosoya, Masaki
; Fujii, Ryo
; Fukusumi, Shoji
; Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,971B
; FILING DATE: 06-Feb-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03821
; FILING DATE: 28-DEC-1996
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 47:
; US-08-776-971-47

Query Match 100.0%; Score 79; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGK 13
Db 12 TPDINPAWYTGK 24

RESULT 15
US-09-421-208-8
; Sequence 8, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

```

STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/421,208
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/105,678
FILING DATE: 26-JUN-1998
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-421-208-8

Query Match 100.0%; Score 79; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGRC 13
Db 12 TPDINPAWYTGRC 24

Search completed: August 12, 2004, 14:52:14
Job time : 8.40698 secs

h